Introduction to Particle Physics

Dong-Sheng Du
Mao-Zhi Yang

World Scientific

Introduction to
Particle Physics

Recommended Titles in Related Topics

Introduction to High Energy Physics: Particle Physics for the Beginner
by Lee G Pondrom
ISBN: 978-981-122-209-2
ISBN: 978-981-122-301-3 (pbk)

Introduction to Quantum Field Theory and the Standard Model
by Wolfgang Hollik
ISBN: 978-981-124-217-5

Particles in the Early Universe: High-Energy Limit of the Standard Model from the Contraction of Its Gauge Group
by Nikolai A Gromov
ISBN: 978-981-120-972-7

Foundations of Quantum Field Theory
by Klaus D Rothe
ISBN: 978-981-122-192-7
ISBN: 978-981-122-300-6 (pbk)

Hadron Spectroscopy and Structure:
Proceedings of the XVIII International Conference
edited by Feng-Kun Guo and Wei-Hong Liang
ISBN: 978-981-121-919-1

Gribov-90 Memorial Volume: Field Theory, Symmetry, and Related Topics — Proceedings of the Memorial Workshop Devoted to the 90th Birthday of V N Gribov
edited by Yu L Dokshitzer, P Lévai, Á Lukács and J Nyíri
ISBN: 978-981-123-839-0

From the Past to the Future: The Legacy of Lev Lipatov
edited by Joachim Bartels, Victor Fadin, Evgeny Levin, Aharon Levy, Victor Kim and Agustin Sabio-Vera
ISBN: 978-981-123-111-7

New Era for CP Asymmetries: Axions and Rare Decays of Hadrons and Leptons
by Ikaros I Bigi, Giulia Ricciardi and Marco Pallavicini
ISBN: 978-981-3233-07-2

New Physics in b Decays
by Marina Artuso, Gino Isidori and Sheldon Stone
ISBN: 978-981-125-129-0

Introduction to Particle Physics

Dong-Sheng Du
Chinese Academy of Sciences, China
Mao-Zhi Yang
Nankai University, China

World Scientific

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI • TOKYO

Published by

World Scientific Publishing Co. Pte. Ltd.
5 Toh Tuck Link, Singapore 596224
USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601
UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

Library of Congress Control Number: 2022013720

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

《粒子物理导论》
Originally published in Chinese by China Science Publishing & Media Ltd.
Copyright © China Science Publishing & Media Ltd. 2015

INTRODUCTION TO PARTICLE PHYSICS

Copyright © 2023 by World Scientific Publishing Co. Pte. Ltd.

All rights reserved. This book, or parts thereof, may not be reproduced in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage and retrieval system now known or to be invented, without written permission from the publisher.

For photocopying of material in this volume, please pay a copying fee through the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, USA. In this case permission to photocopy is not required from the publisher.

ISBN 978-981-125-945-6 (hardcover)
ISBN 978-981-125-946-3 (ebook for institutions)
ISBN 978-981-125-947-0 (ebook for individuals)

For any available supplementary material, please visit
https://www.worldscientific.com/worldscibooks/10.1142/12922#t=suppl

Typeset by Stallion Press
Email: enquiries@stallionpress.com

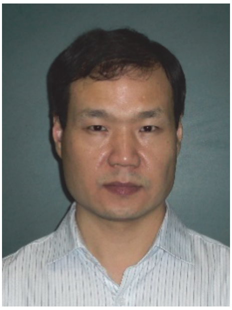
Printed in Singapore

About the Authors

Dong-Sheng Du is a Professor of Physics. He earned his BSc in physics at Peking University in 1964, PhD degree (equivalent) for theoretical physics in 1968 at the Institute of Atomic Physics of Chinese Academy of Sciences. Since 1985, he has been Professor and Adviser of the PhD program at the Institute of High Energy Physics (IHEP). He worked at CERN as Research Associate from 1978 to 1979 and Research Member at Princeton Institute for Advanced Study from 1986 to 1987. He has been the Director of the Theory Division at IHEP, the Chair of Academic Degree Committee and Academic Committee at IHEP, Chairman of the Physics Committee of the National Laboratory of Electron–Positron Collider (BEPC) at IHEP.

His research interests include the Straton model of hadrons, grand unified theories, heavy flavor physics and CP violation, QCD and phenomenology of the standard model, etc. His work on the prediction for the top quark mass lower bound using B-Bbar mixing data has important influence on the top quark search.

Prof. Du founded the popular science magazine “High Energy Physics” (later renamed as “Knowledge of Modern Physics”) and wrote many popular science papers on high energy physics which popularize high energy physics in China.



Mao-Zhi Yang is a Professor of Physics at Nankai University, since 2007. He earned his BSc degree at the Department of Physics, Nankai University, in 1992, and PhD degree at the Institute of High Energy Physics, Chinese Academy of Sciences, in 1998, specializing in theoretical particle physics. From October 1998 to October 2000, he was a postdoctoral fellow at Hiroshima University, Japan, supported by the Japan Society for the Promotion of Science (JSPS). He was an associate professor at Institute of High Energy Physics, Chinese Academy of Sciences, from October 2000 to September 2007. Prof. Yang's research interests include heavy flavor physics, CP asymmetry and perturbative QCD. His research work extends over B physics, charm physics, factorization in QCD, perturbation theory, QCD sum rule, etc.

Preface to the Second Edition

The second edition of the book *Introduction to Particle Physics* has now met with the readers. In the second edition, in addition to correct all the typos and inappropriate points, we have added some new important contents. Among them, we have added: Sec. 5.10 of "Quark Confinement" in Chapter 5; Sec. 9.12 of "The Strong CP Problem" in Chapter 9. We also added Chapter 12 of "Particle Physics and Cosmology". The original Chapter 12 "Epilogue" was moved to Chapter 13. In addition, considering the needs of the beginners we have added Appendix I for the computation of the scattering cross sections for some important physical processes. In Appendix II, we have given the physical and astrophysical constants for the readers' convenience in their practical calculations.

The authors thank Professors Xiaojun Bi, Chaoguang Huang, Zhizhong Xing, Shun Zhou of High Energy Physics Institute and Professor Xin Zhang of Northeast University for many valuable discussions in preparing Chapter 12. We thank, especially, Zhizhong Xing and Shun Zhou for reading the manuscript and their suggestions for Chapter 12. These suggestions that were made in Chapter 12 are more accurate and complete. We also thank the National Natural Science Foundation of China for the support under the grant numbers 11375088, 11375207.

February 2021 The Authors

This page intentionally left blank

Preface to the First Edition

Particle physics is the very important field for investigating the microstructure of matter. Its progress has strong influence not only to all branches of physics, e.g. condensed matter and solid-state physics, cosmology, astrophysics, but also to every field of natural sciences and other science branches, including medical science, agriculture and forestry. This book is designated as an introductory textbook for senior undergraduates, postgraduates, and young researchers, which lays foundations for their theoretical and experimental researches.

The authors have taught particle physics courses for many years in the Graduate School of Chinese Academy of Sciences. This book was written based on the lecture notes given by them there. Since both the experiments and theories were developed very quickly, it is not possible to include all the new advances in one book. But we will do our best, and whenever possible we shall put in new developments and new results.

To the best knowledge of the authors, and due to page limitations, there must be some shortcomings to be improved in the near future. Any comments and criticisms are welcome.

This book was supported in part by the National Natural Science Foundation of China under grant numbers 11375088, 11375207.

July 2014

Dongsheng Du
Institute of High energy Physics,
Chinese Academy of Sciences
Maozhi Yang
Nankai University

This page intentionally left blank

Contents

About the Authors v

Preface to the Second Edition vii

Preface to the First Edition ix

Chapter 1. Overview **1**

References . . . 7

Chapter 2. Symmetry and Conservation Law **9**

2.1 Introduction . . . 9
2.2 The Action, Equation of Motion and Conserved Quantities . . . 10
2.3 Parity Transformation . . . 14
  2.3.1 Klein–Gordon field $\varphi(x)$ . . . 15
  2.3.2 Parity transformation of fermion field . . . 16
  2.3.3 Parity transformation of vector field . . . 18
2.4 Charge Conjugation . . . 19
  2.4.1 Charge conjugation of scalar field . . . 19
  2.4.2 Charge conjugation of fermion field $\psi(x)$ . . . 19
  2.4.3 Charge conjugation of vector field $A_\mu(x)$ . . . 22
2.5 Application: Furry Theorem — A Particle of Spin 1 Cannot Decay to 2γ . . . 22
  2.5.1 Furry theorem . . . 22
  2.5.2 Particles of spin 1 cannot decay into two photons . . . 23

2.6 Time Reversal 25
   2.6.1 Time reversal of Klein–Gordon field $\varphi(x)$ 29
   2.6.2 Time reversal of fermion field 29
   2.6.3 Time reversal of electromagnetic field $A_\mu(x)$ . . . 30
2.7 CPT Theorem 33
References 34

Chapter 3. The Classification and Properties of Particles: Lepton and Hadrons 37

3.1 Four Types of Interactions 37
3.2 Lepton and Lepton Number Conservation 40
   3.2.1 Electron, muon and neutrino 40
   3.2.2 τ and its neutrino ν_τ 43
   3.2.3 Helicity of neutrinos 44
   3.2.4 Lepton number conservation 46
3.3 Hadrons: Conservation of Baryon Numbers 47
   3.3.1 π meson 48
   3.3.2 Nucleon, antinucleon and baryon number conservation 52
   3.3.3 Strange particles 53
   3.3.4 Resonance 57
3.4 Scattering Cross Section, Particle Lifetime and Decay Width 63
   3.4.1 Scattering cross section 63
   3.4.2 Particle lifetime and decay width 67
3.5 Kinematics of Particle Decay 68
   3.5.1 Two-body decay 68
   3.5.2 Three-body decay, Dalitz diagram 69
References 73

Chapter 4. Isospin and G Parity 75

4.1 Isospin 75
   4.1.1 The concept of isospin 75
   4.1.2 Isospin transformation 78
   4.1.3 The law of isospin conservation in strong interaction 83
   4.1.4 The isospin of mesons and baryons 84

4.1.5 An example of isospin analysis of physical processes 88
4.2 Exchange Symmetry: Generalized Identity Principle . . . 90
4.3 Isospin Violation 92
4.4 *G* Parity 93
References 97

Chapter 5. Quark Model of Hadrons 99

5.1 Mathematical Basics 100
5.1.1 Decomposition of SU(n) group representation product, Young tableau 100
5.1.2 Tensor analysis of the SU(3) group 104
5.1.3 SU(3) group generators and Casmir operators 107
5.2 SU(3) Quark Model, the SU(3) Flavor Wave Functions for Mesons and Baryons 109
5.2.1 The flavor wave functions for the pseudoscalar meson octet and singlet 111
5.2.2 Flavor wave functions for the vector meson octet and singlet 114
5.2.3 Flavor wave functions for baryon octet and decuplet 115
5.3 Color Degree of Freedom 122
5.3.1 The relationship of baryon spin and statistics . . . 123
5.3.2 $\pi^0 \to \gamma\gamma$ 123
5.3.3 Measurement of R value in $\mathbf{e}^+\mathbf{e}^-$ annihilation process 124
5.4 Mass Formula of Hadrons 127
5.5 Mixing of Meson Singlet and Octet 129
5.6 OZI Rule 134
5.7 SU(6) Symmetry 135
5.8 Orbital Excitation States and Radial Excitation States, Multi-quark States and Exotic States 137
5.8.1 Orbital and radial excitation states 137
5.8.2 Multi-quark states and exotic states 138
5.9 The Discovery of c, b, t Quarks 139
5.10 Quark Confinement 140
References 141

Chapter 6. Electromagnetic Interaction 143

6.1 QED and Its Feynman Rule 144
6.2 Møller Scattering 149
6.3 Bhabha Scattering 154
6.4 The Electromagnetic Form Factors of Nucleons 156
  6.4.1 Electron–proton elastic scattering assuming proton being a point particle 157
  6.4.2 Elastic scattering of electron–proton 159
6.5 Inelastic Scattering of Electron–Proton 166
  6.5.1 Structure function of inelastic scattering 166
  6.5.2 The structure function applied to elastic scattering 169
6.6 The Parton Model for Nucleons 171
6.7 The Unification of Parton and Quark Model 174
References 177

Chapter 7. Weak Interactions 179

7.1 Looking Back to History 179
7.2 Classification of Weak Decays 181
7.3 Nuclear β Decay 183
7.4 The Discovery of Parity Violation 185
  7.4.1 τ–θ Puzzle 185
  7.4.2 Parity violation in β decays of cobalt 60 nuclei 186
7.5 The V–A Theory of Weak Interactions 187
  7.5.1 Pure leptonic decay 187
  7.5.2 Semi-leptonic decay 189
7.6 Cabibbo Theory and GIM Mechanism 191
7.7 Kobayashi–Maskawa Model 194
7.8 The Limitation of the Four Fermion Point-like Interactions and the Intermediate Vector Bosons 196
7.9 Conservation of Vector Current 198
7.10 Chiral Symmetry Breaking and PCAC 201
7.11 Mixing of Neutral Mesons and CP Violation 205
  7.11.1 Quantum mechanical description 205
  7.11.2 K^0–$\bar{\mathrm{K}}^0$ mixing and CP violation 212
  7.11.3 B^0–$\bar{\mathrm{B}}^0$ mixing and CP violation 224
  7.11.4 D^0–$\bar{\mathrm{D}}^0$ mixing and CP violation 238
References 243

Chapter 8. Gauge Theory of Electroweak Unification 247

8.1 The Higgs Mechanism 248
8.2 Yang–Mills Theory 253
8.3 Glashow–Weinberg–Salam Electroweak Unification Model 259
8.4 The Vacuum Spontaneous Symmetry Breaking: The Higgs Mechanism 268
  8.4.1 The kinetic energy, mass and self-interaction term of scalar field 272
  8.4.2 The mass term of gauge field 274
  8.4.3 The interaction term of the scalar and gauge fields 278
8.5 The Gain of the Fermion Mass: Yukawa Coupling 279
8.6 The Interaction of Fermion and Gauge Field II 285
8.7 The Interaction of Gauge Fields 288
8.8 Feynman Rule in the Renormalization Gauge (R_ξ Gauge) 294
  8.8.1 For pure scalar field 295
  8.8.2 For the interaction of scalar and gauge fields . . . 295
  8.8.3 For ghost fields 298
  8.8.4 The kinetic energy term of the Lagrangian in R_ξ gauge 302
  8.8.5 The interaction terms of scalar and fermions . . . 303
8.9 The Discovery of Higgs Boson 310
References 314

Chapter 9. Theory of Strong Interaction: QCD 315

9.1 The SU(3) Gauge Symmetry of Color 316
9.2 The Quantization of Gauge Field and Fermion Field . . . 320
  9.2.1 The path integral quantization 320
  9.2.2 Reduction formula 328
  9.2.3 The quantization of gauge field 331
  9.2.4 The quantization of fermion field and Grassmann algebra 336
9.3 The Effective Lagrangian of QCD and Perturbation Theory 342
  9.3.1 The effective Lagrangian of QCD 342
  9.3.2 Perturbation expansion 344

9.4 Divergence and Regularization of Loop Corrections 356
9.5 The Renormalization of Divergence 367
  9.5.1 The superficial degree of divergence 368
  9.5.2 The renormalizability of a theory 371
  9.5.3 The renormalization of QCD 375
  9.5.4 One-loop result of the renormalization constants 385
9.6 BRST Symmetry and the Generalized Ward–Takahashi Identity 391
  9.6.1 BRST symmetry 391
  9.6.2 The generalized Ward–Takahashi identity (Slavnov–Taylor identity) 395
9.7 The Renormalization Group Equation and the Solution 401
  9.7.1 The renormalization group 401
  9.7.2 The renormalization group equations 409
  9.7.3 The solution of the renormalization group equation 413
9.8 Asymptotic Freedom 416
9.9 Structure Function, Parton Distribution Function and the Application in Perturbative QCD 424
  9.9.1 Structural function of deep inelastic scattering process 424
  9.9.2 The prediction of QCD to the variation of structure function with Q^2 427
  9.9.3 Examples for application 429
9.10 A Brief Introduction to Some Bound States in QCD . . . 432
  9.10.1 e^+e^- collision, charmonium and bottomonium 432
  9.10.2 Glueball and hybrid state 434
9.11 Nonperturbative QCD and Lattice Gauge Theory 437
9.12 The Strong CP Problem 439
References 451

Chapter 10. Neutrino Oscillation 455

10.1 Experimental Evidence for Neutrino Oscillation 457
  10.1.1 Atmospheric neutrino oscillation 457
  10.1.2 Accelerator neutrino oscillation 459

10.1.3 Solar neutrino oscillation 460
10.1.4 Reactor neutrino oscillation 462
10.2 Theoretical Description of Neutrino Oscillation 464
10.2.1 Neutrino oscillation in vacuum 464
10.2.2 Neutrino oscillation in matter 468
10.3 Dirac Neutrino and Majorana Neutrino 472
10.3.1 Dirac neutrino 472
10.3.2 Majorana neutrino 474
10.4 Neutrino Mass and See-Saw Mechanism 476
10.5 Parameterization of Neutrino Mixing Matrix 480
10.6 CP Violation in Neutrino Oscillation 481
10.7 Neutrino Mass Hierarchy Problem 482
10.8 Perspectives 484
References 486

Chapter 11. Beyond Standard Model 489

11.1 Grand Unified Theories for Strong, Electromagnetic and Weak Interactions 490
11.1.1 SU(5) grand unified theory 490
11.1.2 SO($2N$) grand unified models 509
11.1.3 Flavor unification 509
11.2 Supersymmetric Models 510
11.2.1 Superspace and superfields 513
11.2.2 Global supersymmetry and local supersymmetry (Supergravity) 514
11.2.3 Global supersymmetric theory and minimal supersymmetric standard model 515
11.3 Superstring Theory and Brane Theory 526
References 527

Chapter 12. Particle Physics and Cosmology 529

12.1 Basics of the Big Bang ΛCDM Model 530
12.1.1 Hubble's law 531
12.1.2 The Robertson–Walker metric and the Friedmann equation 532
12.1.3 Definitions of some cosmological parameters . . . 534
12.1.4 The redshift and its relation with the physical distance 536

12.1.5 Some important solutions to the standard ΛCDM model 538
12.1.6 The age of the Universe 540
12.2 Hot Big Bang and Radiation in the Early Universe 542
12.3 Neutrino Decoupling and Cosmic Neutrino Background 548
12.4 Big Bang Nucleosynthesis and Abundances of Primordial Light Elements 551
12.5 Cosmological Baryon–Antibaryon Asymmetry 555
12.6 Dark Matter 558
12.7 Dark Energy 561
12.8 Cosmic Microwave Background 565
12.8.1 Transition of the Universe to its matter-dominated epoch 565
12.8.2 Formation of the CMB 567
12.8.3 Anisotropies of the CMB 568
12.9 Three Cosmic Problems and the Inflation Mechanism . . . 572
12.9.1 Three problems in cosmology 572
12.9.2 The inflation mechanism 574
References 578

Chapter 13. Epilogue **581**

Appendix I. Calculation of Some Cross Sections **585**

Appendix II. Table of Physical Constants **609**

Index 617

Chapter 1

Overview

Particle physics is the science of particle interactions and mutual transformations. The main aim is to find a simple as well as universal law for explaining the phenomena of all kinds of particle interactions and mutual transformations. For example, the Maxwell equation can give a unified explanation for all electromagnetic phenomena, and the unified theory of the electromagnetic and weak interactions can provide unified explanation for all phenomena of the electroweak interactions, etc. Of course, till now we still do not find a basic theory to give a unified explanation for all particle phenomena. Although the superstring theory (in 10-dimensional space) and membrane theory (in 11-dimensional space) can unify all the four kinds of interactions, but these two theories meet serious difficulties when they compact into realistic 4-dimensional (3+1-dimensional) space–time. This problem needs further investigation.

Particle physics is also called "Theory of elementary particles" or "high-energy physics". The last two appellations are inherited from the history. Elementary Particles are the generic term for proton, neutron, mesons, hyperons, electron, photon, neutrino, etc. The term "elementary" means that no further division can be made! In history, people thought, the above elementary particles were the smallest basic building blocks of matter; they could not be divided further. The science which investigates the law of interactions and mutual transformations of elementary particle is called "physics of elementary particles". The appellation of "high energy physics" comes from the fact that investigating elementary particle properties needs higher energy than that for studying atomic physics and nuclear physics. In fact, to break an atom or to kick out an electron from an atom needs to pay about 10 eV energy. But to kick out a nucleon (proton or neutron) needs to pay about 10 MeV energy which is a million times higher than that for the case of atoms. How much energy we should pay for breaking an "elementary" particle, i.e. a proton or a neutron? Now we know at least we need to pay for several tens or several hundreds times more of energy. We can see from above discussion that the division of "high energy" and "low energy" does not have an accurate energy dividing line. So the appellation for "high energy" and "low energy" is not exact or not appropriate. Today, we already know that elementary particles are not "elementary". They can be further divided or broken apart. From the modern concept of view, the name "elementary particle physics" is unscientific. So the appellation "particle physics" is more appropriate. Of course, it is very difficult to give "particle physics" a rigorous definition. No matter what definition it is, the core idea is that "particle physics" is a science which searches for the substructure of matter and the law of the basic interactions of nature. Thus "particle physics" is the scientific frontier for investigating microscopic structure of matter and the corresponding rules. In 1957, IUPAP gave a definition for particle physics, that is, "particle physics" is the science for investigating the properties of the basic building blocks of matter and their interactions. But this definition needs to add new content. Why? Because very important advances have happened in recent years in astrophysics and cosmology, these new advances and developments connect particle physics with astrophysics and cosmology. Observations of astronomers revealed that the redshift of the light delivered by supernova is larger than that predicted by the empirical Hubble law. This means that the supernova is escaping away from us with acceleration, that is, our universe is expanding with acceleration. So the universe needs to have negative pressure. One possible choice is to assume that the universe is fully filled by dark energy which provides the negative pressure. Astronomic observations tell us that the dark energy occupies 69% of the total energy density of the universe, the dark matter 26%, the baryonic matter 5% in which only 1% are visible objects (galaxies, dust, etc.). What is the dark energy? Is it a scalar field which fulfilled the entire universe or the Einstein cosmological constant? We do not have any conclusion till now. At present, dark matter and dark energy are both the research topics at the frontier of particle physics. In addition, the evolutions of the early universe after the Big Bang and the formation of the baryon–antibaryon asymmetry, and from nothing to everything of the universe, the nucleosynthesis, the formation of galaxies and the cosmic microwave background, etc., need to be explained by using particle physics. So, two interdisciplines, i.e., "particle cosmology" and "particle astrophysics", are developed in recent years. In addition, the development of neutrino physics combines with astrophysics to form an another new interdiscipline "neutrino astrophysics".

According to the above new developments in recent years, American "High Energy Physics Advisory Panel — HEPAP" proposed a definition for particle physics [1], that is "High-energy physics is the science of interactions", i.e. "Interactions, the science of matter, space and time". Considering the lack of knowledge for dark energy, we can give a new definition for particle physics or high-energy physics: "Particle physics (high-energy physics) is the science of matter, energy, space and time". This new definition is more appropriate, more comprehensive, and has the characteristics of the times. The content of this book totally reflects the profound connotation of this new definition.

The first "elementary particle" discovered is electron [2]. (In 1897, J. J. Thomson firstly measured the ratio of electric charge over mass e/m of a negatively charged particle in the cathode ray. In 1899, he measured also its electric charge.) The photon γ as a particle was firstly proposed by Einstein in 1905 [3]. (Einstein tried to explain photoelectric effect.) But Einstein's explanation was recognized by physics community until more than 20 years later. E. Rutherford in 1919 discovered proton using α particle to bombard nitrogen [4]. In 1932, J. Chadwick repeated the experiment done by the Curies in 1931 carefully (α particle bombarding beryllium) and established the existence of neutron for the first time [5]. In the same year (1932), C.D. Anderson discovered the positron in the cosmic-ray experiment using cloud chamber [6]. (At that time, Anderson did not know that Dirac theory already predicted the existence of positron.) The famous Chinese scientist Chong-yao Chao already found annihilation radiation (in fact this is $e^+e^- \to 2\gamma$) before the discovery of positron; he had come to the entrance for discovering the positron. Later, a series of particles were found in cosmic-ray experiments: μ [7], π meson [8], K meson, etc. [9]. Until the 1950s of the 20th century, because the emerging of artificial particle accelerators a large number of particles and resonances were discovered, e.g., Λ, Σ, Ξ, Δ_{33}, $\bar{\rm p}$, $\bar{\rm n}$. From then on the research on "elementary particles" turned round the situation of "depending on Heaven for food", people could do what they want by building all kinds of accelerators. In later years, the number of "elementary particles" increased to about a hundred. Comparing with periodic table of the elements, we can naturally imagine that about one hundred "elementary" particles should not be all elementary. They must have internal structure. The existence of the hundred elementary particles and resonances is the foundation for building the quark model of hadrons.

In order to give classification to "elementary particles", we come back to the basic interactions in Nature. Today, we know that there are four kinds of the basic interactions in Nature: gravitation, weak interactions, electromagnetic interactions, and strong interactions (nuclear force). Their strengths can be characterized by the following dimensionless constants:

- gravitation: $\frac{G_{\mathrm{N}}m_{\mathrm{N}}^2}{\hbar c} \sim 0.6 \times 10^{-38}$, force range ∞;
- weak interaction: $G_{\mathrm{F}}m_{\mathrm{N}}^2\frac{c}{\hbar^3} \sim 10^{-5}$, force range 10^{-16} cm;
- electromagnetic interaction: $\frac{e^2}{4\pi\hbar c} \sim 10^{-2}$, force range ∞;
- strong interaction: $\frac{g_{\mathrm{s}}^2}{4\pi} \geq 1$, force range 10^{-13} cm.

The above constants are the values at the scale $m_{\mathrm{N}} \sim 1\mathrm{GeV}$, In fact, the values of the above constants depend on the energy scale. For instance, at 1 GeV, for electromagnetic interactions $\alpha = \frac{e^2}{4\pi\hbar c} \sim \frac{1}{137}$, this coupling constant increases to $\alpha \approx \frac{1}{128}$ at the scale of 100 GeV. While for the strong interactions $\alpha_{\mathrm{s}} = \frac{g_{\mathrm{s}}^2}{4\pi}$ turns to zero gradually (asymptotic freedom) when the scale tends to ∞. The gravitation interaction is very weak between elementary particles, so in most cases of observation, gravitation can be neglected.

Now we can classify the elementary particles after knowing the strengths of the four interactions. All particles participating strong interactions (of course, they can participate other interactions at the same time) are called "hadrons". Hadrons include mesons and baryons. Mesons include π, K, ϱ, K*, etc. Baryons include proton, neutron, and Λ, Σ, Δ, etc. Baryons' spins are either 1/2 or 3/2. They are called "baryons" just because their masses are heavier than mesons. Fermions which do not participate strong interactions and their masses are lighter than mesons are called "leptons". But there is one exception, i.e. the heavy lepton τ. Its mass is almost twice of the proton's, but its physical properties are the same as e and μ, so it is called "heavy lepton". In one word, p, n, π, Λ, Σ, Δ, etc. are hadrons, e, μ, τ, ν_{e}, ν_μ, ν_τ, etc. are leptons.

As mentioned above, about one hundred hadrons are not possibly all "elementary", they must have internal structure, or substructure.

In 1956, Hofstadter [10] (Stanford University) firstly observed that the electric charge and magnetic moment of the proton are distributed in a limited space, while the neutron's magnetic moment is also distributed in a limited space. So the internal structures of the proton and neutron were verified.

In 1964, M. Gell-Mann [11] proposed the quark model of hadrons, in which, all hadrons are composed from three kinds of quarks u, d, s and their anti-quarks. At that time, the classification of hadrons using SU(3), SU(6) symmetries became fashionable. Gell-Mann got Nobel Price in 1969 because of this model.

In 1967 [12] at the Stanford Liner Accelerator Center (SLAC) the scaling property of the nucleon structure functions was observed in the electron–proton deep inelastic scattering experiments. This result meant that there exist inside protons a large number of spin-1/2 point-like charges (called partons) moving almost freely. This hastened the emerging of the parton model [13], and established the asymptotic freedom of the strong interactions. (Strong interactions become weaker at higher energies and turn to zero when energy turns to infinity.) Quantum chromo dynamics (QCD) appeared in 1973 [14] unifies the low-energy picture of the substructure of the proton (consist of three valence quarks uud) and high-energy picture (consist of three valence quarks uud and infinite number of $q\bar{q}$ pairs and gluons), but the incoming electron cannot see the gluons because the charge neutral gluons cannot interact with the electron.

In 1974, Ting and Richter [15] firstly observed J/Ψ particle (consist of $c\bar{c}$) at BNL (Brookhaven National Laboratory) and SLAC (Stanford Linear Accelerator Center) respectively, so the charm particle was discovered, and the existence of the quark c with the fourth “flavor” was established.

In 1977, L. Lederman [16] observed a new particle Υ (9.5 GeV) in the experiments of 400 GeV proton bombarding nuclear target at Fermi laboratory. This new particle is actually a bound state of another new quark b and its anti-quark $\bar{b}$ ($b\bar{b}$), so the existence of the fifth “flavored” quark b was established.

In 1994, CDF collaboration [17] at Fermi laboratory discovered the top quark (t-quark). Its mass is about 180 times of the proton mass.

Until now, we totally discovered, u, d, c, s, t, b, six quarks with different flavor. Each “flavor” has three different “color”. The quarks with same flavor but different color have same mass and spin. The existence of the new quantum number “color” has been verified by the measurements of $\pi^0 \to 2\gamma$ and $R = \frac{\sigma\left(e^+e^- \to \text{hadrons}\right)}{\sigma\left(e^+e^- \to \mu^+\mu^-\right)}$ (see Chapter 5: Quark Model of Hadrons).

The $Z^0 \to \nu\bar{\nu}$ experiments at LEP (Large Electron Positron colliding machine in Geneva) revealed that there are only three types of neutrinos with mass less than $m_Z/2 \sim 45.5\text{GeV}$, i.e. ν_e, ν_μ, ν_τ. The common believe now in physics community is that there are only three generations(or families) of quarks and leptons, i.e.

$$\begin{array}{ccc} \text{first generation} & \text{second generation} & \text{third generation} \\ \begin{pmatrix} \mathrm{u} \\ \mathrm{d} \end{pmatrix} & \begin{pmatrix} \mathrm{c} \\ \mathrm{s} \end{pmatrix} & \begin{pmatrix} \mathrm{t} \\ \mathrm{b} \end{pmatrix} \\ \begin{pmatrix} \nu_{\mathrm{e}} \\ \mathrm{e} \end{pmatrix} & \begin{pmatrix} \nu_{\mu} \\ \mu \end{pmatrix} & \begin{pmatrix} \nu_{\tau} \\ \tau \end{pmatrix} \end{array}$$

Why are there only three generations? We still do not know. Further investigations are needed.

For the development of theoretical particle physics, we have gone through several important breakthroughs in the history.

In 1905, A. Einstein [18] proposed the special relativity theory. This made our knowledge about Nature increased from Newtonian mechanics to relativistic mechanics.

In 1923, L. de Broglie [19] proposed the concept of matter wave, and in 1926 E. Schrödinger [20] set up wave mechanics (non-relativistic quantum mechanics) based on the mater wave concept. This made a revolutionary breakthrough about our knowledge on the law of the microscopic world.

In 1927, P. A. M. Dirac [21] firstly completed the quantization of the electromagnetic fields.

In 1928, Dirac proposed Dirac equation [22] based on the combination of the non-relativistic quantum mechanics and special relativity. So the concept of antiparticle appeared. This made our knowledge on microscopic world to go forward one large step.

In 1929, Heisenberg and Pauli [23] presented the general form of quantum field theory.

In 1943∼1949, covariant field theories were set up [24].

In 1954, Yang and Mills [25] proposed non-Abelian gauge fields, but they did not know how to gave particle masses. The appearance of quantum field theory and non-Abelian gauge fields set up a solid foundation for building the correct theories for describing the mutual interactions of elementary particles.

In 1967, the unified theory of electroweak interactions was proposed [26], this theory is builded on the base of the $\mathrm{SU}(2) \times \mathrm{U}(1)$ non-Abelian gauge field theory, and the problem how to give particle masses was solved by introducing the Higgs mechanism. This unified theory went through all the electroweak experiments and was verified that it is a correct theory. This is the one of the greatest achievements of particle physics.

In 1973, the quantum chromo dynamics (QCD) was builded on the bases of SU(3) non-Abelian gauge theory [13]. QCD solves theoretically the problem of strong interactions, and its high-energy behavior is proved to be correct by large amount of experiments.

In 1974, a unified theory for strong, electromagnetic, and weak interactions appeared, that is the SU(5) grand unified model [27]. But this model was not supported by experiments.

In the 1980s of the 20th century, string theory [28] and membrane theory [29] were proposed successively. These two theories can unify the four kinds of strong, weak, electromagnetic, and gravitation interactions. But because these two theories are builded on high-dimensional space, when they are compacted into the realistic four dimensional space–time, they meet serious difficulties. These problems need further investigation.

In the following, we summarize graphically the history and present status of the research on microscopic structure of matter. For every microscopic level, we specify the corresponding physics in its followed brackets, and the arrows express going into next deeper level: *Atoms (atomic physics, quantum mechanics)* → *nuclei (nuclear physics, quantum mechanics)* → *"elementary particles", proton, neutron, electron, etc.(particle physics, high-energy physics, quantum field theories, string theories, etc.)* → *quarks and leptons (subparticle physics? super high-energy physics?)* → *preon (? physics)* → ⋯ → *? (? ? physics).*

In this chapter, we reviewed briefly the history and present status for the development of particle physics and answered what is particle physics and its subjects investigated. In the following chapters, we will elucidate the basic content of particle physics and lay research foundation for the readers.

References

1. Report of HEPAP[a].
2. Thomson J. *J. Phil. Mag.*, 1897, 44: 293.
3. Einstein A. *Ann. Phys.*, 1905, 17: 132.
4. Rutherford E. *Phil. Mag.*, 1919, 37: 581.
5. Chadwick J. *Nature*, 1932, 129: 312.
6. Anderson C D. *Phys. Rev.*, 1932, 41: 405.
7. Anderson C D, Neddermeyer S H. *Phys. Rev.*, 1936, 50: 273.
8. Powell C F. *Nature*, 1947, 160: 453.

---

[a] High Energy Physics Advisory Panel, USA, 2001. We thank Professor Bing-Lin Young of Iowa State University for introducing this to us.

9. Rochester G D, Butler C C. *Nature*, 1947, 160: 855.
10. McAllister R W, Hofstadter R. *Phys. Rev.*, 1956, 102: 851; Chambers E E, Hofstadter R. *Phys. Rev.*, 1956, 103: 1454; McAllister R W. *Phys. Rev.*, 1956, 104:1494.
11. Gell-Mann M. *Phys. Lett.*, 1964, 8: 214.
12. Panofsky W K H. *Proc. 14th Int. Conf. High Energy Phys.*, Vienna, Austria, 1968: 23; Friedman J I, Kendall H W. *Ann. Rev. Nucl. Sci.*, 1972, 22: 203.
13. Bjorken J D, Paschos E A. *Phys. Rev.*, 1969, 185: 1975; Feynman R P. *Phys. Rev. Lett.*, 1969, 23: 1415.
14. Politzer H D. *Phys. Rev. Lett.*, 1973, 30: 1346; Gross D J, Wilczek F. *Phys. Rev. Lett.*, 1973, 30: 1343.
15. Aubert J J, *et al. Phys. Rev. Lett.*, 1974, 33: 1404; Augustin J E, *et al. Phys. Rev. Lett.*, 1974, 33: 1406.
16. Herb S W, *et al. Phys. Rev. Lett.*, 1977, 39: 252; Innes W R, *et al. Phys. Rev. Lett.*, 1977, 39: 1240.
17. CDF Collab. *Phys. Rev. Lett.*, 1995, 74: 2626.
18. Einstein A. *Ann. Phys.*, 1905, 17 (4): 891.
19. de Broglie L. *Nature*, 1923, 112: 540; de Broglie L. *Ann. Phys.*, 1925, 3: 22 .
20. Schrödinger E. *Ann. Phys.*, 1926, 79: 361, 489, 734; 1926, 80: 437; 1926, 81: 109 .
21. Dirac P A M. *Proc. Roy. Soc.*, 1927, A114: 243, 710.
22. Dirac P A M. *Proc. Roy. Soc.*, 1928, A117: 610; ibid, 1928, A118: 341.
23. Heisenberg W, Pauli W. *Zeits. Phys.*, 1929, 56: 1; 1929, 59: 168.
24. Schwinger J. *Phys.*, 1948, 74: 1439; Tomonaga S, *et al. Prog. Theor. Phys.*, 1949, 4: 47,121; Feynman R P. *Phys. Rev.*, 1949, 76: 749, 769.
25. Yang C N, Mills R L. *Phys. Rev.*, 1954, 96: 191.
26. Weinberg S. *Phys. Rev. Lett.*, 1967, 19: 1264; Glashow S. L. *Nucl. Phys.*, 1961, 22: 579; Salam A. in *Elementary Particle Physics* (*Nobel Symp* No. 8, 1968), ed. Svartholm N, Almqvist and Wilsell.
27. Georgi H, Glashow S L. *Phys. Rev. Lett.*, 1974, 32: 438.
28. Polchinski J. hep-th/9411028.
29. West P C. hep-th/9811101.

Chapter 2

Symmetry and Conservation Law

2.1. Introduction

The Lagrangian form of classical mechanics, namely analytical mechanics, is a highly developed form of Newtonian mechanics, where only energy, potential, coordinate, velocity or momentum are involved. Force is not a frequently considered concept in analytical mechanics. Lagrangian and Hamiltonian system of analytical mechanics are more appropriate to be used to construct field theory, where particle interactions are described. Classic field equation can be derived from the action obtained from the Lagrangian of fields. Symmetry is an important concept in field theory. It reflects the invariance of the interaction under some specific transformation. Symmetry can be inferred from the action of the system. It can be proved that there must be conserved quantity and conserved current if there is any symmetry in the system.

Symmetries play an important role in the study of modern physics. Some symmetries are strict. Some are approximate. Some symmetries are true under certain conditions and broken in general. Broken symmetry are also important in particle physics.

There are four types of symmetries found in physics till now [1]. They are symmetry of permutation, continuous symmetry of space and time, discrete symmetry and unitary symmetry. Permutation symmetry denotes the symmetry of the system under interchange within states of two or more identical particles; continuous symmetry means the invariance of the physical system under continuous transformation of space and time, such as, space rotation invariance, spacial translation invariance and time displacement invariance, etc.; unitary symmetry is an intrinsic symmetry within the particle states, such as, the U(1) symmetry, which is related to the phase transition of particle fields. Electric change and baryon number conservations are all consequences of U(1) symmetry. SU(2) symmetry of multiplet fields and SU(3) symmetry for hadrons are also unitary system. All these symmetries are gauge symmetries. Gauge symmetries can be further separated into local gauge symmetry and global gauge symmetry. For global gauge symmetry, the standard electroweak theory and quantum chrome dynamics (QCD) are all based on local gauge symmetries.

2.2. The Action, Equation of Motion and Conserved Quantities

Particles are described by the relevant fields in quantum field theory. Particles are created as fields excited. Field will jump back to vacuum state when it is completely de-excited. Physical vacuum is not a pure null state. It is only a state of the lowest-energy level of field.

Consider a classic field $\phi(x)$ with Lagrangian density $\mathcal{L}(\phi, \partial_\mu \phi)$. The action of the field system is

$$S = \int \mathrm{d}^4 x \mathcal{L}(\phi, \partial_\mu \phi) \tag{2.2.1}$$

Consider the condition that the field $\phi(x)$ tends to be zero at the point of space–time to be infinity, i.e., $\phi(x \to \pm\infty) = 0$. When there is an infinitesimal change for the field $\phi(x)$,

$$\phi(x) \to \phi(x) + \delta\phi(x) \tag{2.2.2}$$

The action will change relevantly

$$\begin{aligned}\delta S &= \int \mathrm{d}^4 x \left(\frac{\partial \mathcal{L}}{\partial \phi} \delta\phi + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \delta(\partial_\mu \phi) \right) \\ &= \int \mathrm{d}^4 x \left\{ \left(\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right) \delta\phi + \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \delta\phi \right) \right\}\end{aligned} \tag{2.2.3}$$

The second term in the above equation is an integral of a 4-divergence term, which can be turned to be an surface integral. Therefore, the second term in Eq. (2.2.3) turns to be zero under the condition that the change of the field at the point of infinity is fixed to be zero,

$$\delta S = \int \mathrm{d}^4 x \left(\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right) \delta\phi \tag{2.2.4}$$

According to the principle of least action, the actual path that really happens during the variation of the field shall be that for which the action takes the extreme value. So there is

$$\delta S = 0 \tag{2.2.5}$$

Combining Eqs. (2.2.4) and (2.2.5), the classic equation of motion can be obtained

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} - \frac{\partial \mathcal{L}}{\partial \phi} = 0 \tag{2.2.6}$$

This equation is called Euler–Lagrange equation of motion.

If there is any symmetry for the system, i.e., the system is invariant under some variations of coordinate and the value of the field, then the relevant change of the action is

$$\delta S = \int \left\{ \delta(\mathrm{d}^4x)\mathcal{L} + \mathrm{d}^4x\delta\mathcal{L} \right\} \tag{2.2.7}$$

where δx^μ and $\delta\phi$ are the variations of the space–time coordinate and the value of the field. Up to the order of $O(\delta x)$, the variation of the integral measure is

$$\delta \mathrm{d}^4x = \mathrm{d}^4x' - \mathrm{d}^4x = \mathrm{d}^4x\partial_\mu\delta x^\mu \tag{2.2.8}$$

The variation of the Lagrangian density is

$$\delta\mathcal{L} = \delta x^\mu \partial_\mu \mathcal{L} + \frac{\partial \mathcal{L}}{\partial \phi}\delta_0\phi + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\delta_0\partial_\mu\phi \tag{2.2.9}$$

where the operator δ_0 is defined as $\delta_0\phi(x) = \phi'(x) - \phi(x)$. Here the operator δ_0 only includes the variation of the field $\phi(x)$ at the same space–time coordinate, but does not include the variation caused by the change of the coordinate. Since $\delta_0(\partial_\mu\phi) = \partial_\mu(\delta_0\phi)$, consequently Eq. (2.2.9) becomes

$$\delta\mathcal{L} = \delta x^\mu \partial_\mu \mathcal{L} + \left[\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\right]\delta_0\phi + \partial_\mu\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\delta_0\phi\right) \tag{2.2.10}$$

Considering the classical equation of motion, then

$$\delta\mathcal{L} = \delta x^\mu \partial_\mu \mathcal{L} + \partial_\mu\left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\delta_0\phi\right) \tag{2.2.11}$$

By substituting Eqs. (2.2.8) and (2.2.11) into Eq. (2.2.7), the change of the action will be

$$\begin{aligned}\delta S &= \int \mathrm{d}^4x \left[(\partial_\mu \delta x^\mu)\mathcal{L} + \delta x^\mu \partial_\mu \mathcal{L} + \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\delta_0 \phi\right)\right] \\ &= \int \mathrm{d}^4x \partial_\mu \left[\delta x^\mu \mathcal{L} + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\delta_0 \phi\right] \qquad (2.2.12)\end{aligned}$$

Consider the relation of $\delta_0\phi(x)$ and $\delta\phi(x)$, i.e.

$$\begin{aligned}\delta\phi(x) &\equiv \phi'(x') - \phi(x) \\ &= \phi'(x+\delta x) - \phi(x) \\ &= \phi'(x) + \delta x^\mu \partial_\mu \phi'(x) - \phi(x) \qquad (2.2.13)\end{aligned}$$

The second term in the above equation can be substituted directly by $\phi'(x) \to \phi(x)$, because only contributions of first order of δx and/or $\delta_0\phi$ are considered. $\phi'(x) - \phi(x)$ is at the first order of $O(\delta x)$ and $O(\delta_0\phi)$, and the difference for the second term caused by the substitution of $\phi' \to \phi$ is at the second order of δx and/or $O(\delta_0\phi)$, so we can directly substitute $\phi'(x)$ by $\phi(x)$. Then Eq. (2.2.13) becomes

$$\begin{aligned}\delta\phi(x) &= \phi'(x) + \delta x^\mu \partial_\mu \phi(x) - \phi(x) \\ &= \delta_0\phi(x) + \delta x^\mu \partial_\mu \phi(x) \qquad (2.2.14)\end{aligned}$$

The significance of Eq. (2.2.14) is that the total variation of the field $\phi(x)$ equals to the sum of the variation of the field at the same space–time coordinate and that caused by the change of the coordinate. Now $\delta_0\phi(x)$ can be expressed as

$$\delta_0\phi(x) = \delta\phi(x) - \delta x^\mu \partial_\mu \phi(x) \qquad (2.2.15)$$

Therefore, Eq. (2.2.12) becomes

$$\delta S = \int \mathrm{d}^4x \partial_\mu \left[\left(\mathcal{L} g^\mu_\nu - \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\partial_\nu \phi\right)\delta x^\nu + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)}\delta\phi\right] \qquad (2.2.16)$$

Because the variations of the field and coordinate δx and $\delta\phi$ are caused by any symmetric transformation, they can be expressed by a set of common transformation parameters ω^a (where $a = 1, 2, 3, \ldots, N$),

$$\delta x^\nu = \frac{\delta x^\nu}{\delta \omega^a}\delta\omega^a, \quad \delta\phi = \frac{\delta\phi}{\delta\omega^a}\delta\omega^a \tag{2.2.17}$$

Therefore, the change of the action δS is

$$\delta S = \int \mathrm{d}^4 x \partial_\mu \left[\left(\mathcal{L} g^\mu_\nu - \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\nu\phi\right)\frac{\delta x^\nu}{\delta\omega^a} + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\frac{\delta\phi}{\delta\omega^a}\right]\delta\omega^a \tag{2.2.18}$$

Under the symmetric transformation, the change of the action must be zero, i.e., $\delta S = 0$, for any value of the variation parameters $\delta\omega^a$. Therefore, it can be obtained

$$\partial_\mu \left[\left(\mathcal{L} g^\mu_\nu - \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\nu\phi\right)\frac{\delta x^\nu}{\delta\omega^a} + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\frac{\delta\phi}{\delta\omega^a}\right] = 0 \tag{2.2.19}$$

Equation (2.2.19) is called equation of current conservation. If the system is invariant under some kind of transformation, then the current defined through Eq. (2.2.20)

$$j^\mu_a = -\left(\mathcal{L} g^\mu_\nu - \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\nu\phi\right)\frac{\delta x^\nu}{\delta\omega^a} - \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\frac{\delta\phi}{\delta\omega^a} \tag{2.2.20}$$

is conserved,

$$\partial_\mu j^\mu_a = 0 \tag{2.2.21}$$

This is Noether's theorem in the classic field theory [2]. We can obtain the conservation laws relevant to some specific symmetric transformations according to Noether's theorem (see Table 2.1).

Here, as an example, only about how U(1) symmetry generating conserved vector current in Table 2.1 is briefly discussed. Let us consider the

Table 2.1. The symmetry and the relevant conservation law.

Symmetry	Conservation law
Time translation invariance	Energy conservation
Space translation invariance	Momentum conservation
Space rotation invariance	Angular momentum conservation
Global U(1) symmetry	Vector current conservation

case of Fermion system. The Lagrangian of the system is

$$\mathcal{L} = \mathrm{i}\overline{\Psi}\gamma^\mu \partial_\mu \Psi - m\overline{\Psi}\Psi \tag{2.2.22}$$

U(1) symmetry apparently holds in such a system. The Lagrangian is invariant under the following transformation of the field:

$$\Psi \to \mathrm{e}^{\mathrm{i}\alpha}\Psi, \quad \overline{\Psi} \to \overline{\Psi}\mathrm{e}^{-\mathrm{i}\alpha} \tag{2.2.23}$$

Substituting the field ϕ in Eq. (2.2.20) by the Fermion field ψ, changing the transformation parameter as α, and using $\delta\Psi = \mathrm{i}\alpha\Psi$ and $\delta x = 0$ in Eq. (2.2.20), we can obtain the conserved current

$$j^\mu = \overline{\Psi}\gamma^\mu \Psi, \quad \partial_\mu j^\mu = 0 \tag{2.2.24}$$

For the other continuous symmetries of the classical field such as Lorentz invariance and Poincare invariance, the readers can refer to books on quantum field theory, for example Ref. [3]. No further references are given explicitly here. Several important discrete transformations in particle physics, such as parity transformation, charge conjugate transformation and time-reversal transformation, will be discussed in the subsequent sections.

2.3. Parity Transformation

Parity transformation is denoted by a symbol $\hat{P}$, which is also called space reflection. It changes the space coordinate as $\vec{x} \to -\vec{x}$. A left-hand coordinate system is transformed to a right-hand system under parity transformation. Therefore, parity transformation is also called left–right transformation. If a physical system is invariant under parity transformation, this system is a parity-conserved system. The reason that parity is very important in particle physics is that parity conservation has ever been an important principle for writing the Lagrangian for particle interaction in history.

The symmetry of left and right is an important experience that we obtain from everyday life. Our left and right hands are very similar to each other. The face of human being is symmetric between the left and right parts. These are very interesting phenomena. In addition, a car moving towards left can be transformed to moving towards right in a mirror. Moving towards right can certainly happen in the real world. A fan rotating towards left can rotate towards right if we reverse the moving direction of the fan blades. The nature prefers neither left nor right. Left and right are only a convention of human being. Left and right can actually not be distinguished absolutely. The laws of nature are all left–right symmetric. These experience obtained from macroworld is carelessly extended into microworld. Parity conservation has been used to analyze physical processes of electromagnetic, strong and weak interactions. It was thought that parity is conserved in all interactions, including weak interaction. This led to a serious puzzle.[a] The problem was kept unsolved until parity violation was suggested by Lee and Yang in 1956 [4]. In the following, parity transformation of Klein–Gordon field $\varphi(x)$, fermion field $\psi(x)$ and vector field, such as electromagnetic field, will be discussed.

2.3.1. *Klein–Gordon field $\varphi(x)$*

The field $\varphi(x)$ satisfies Klein–Gordon equation $(\partial_\mu\partial^\mu + m^2)\varphi(x) = 0$. Klein–Gordon equation should be invariant under parity transformation. Then the transformation of the field $\varphi(x)$ is $\varphi(x) \to \pm\varphi(\tilde{x})$, $\tilde{x}^\mu = x_\mu$, where $\pm$ is the intrinsic parity of φ. The sign $+$ is for scalar field, while $-$ for pseudoscalar field. Parity transformation can be expressed in terms of parity transformation operator in Hilbert space $\hat{P}$ as

$$\begin{aligned} \text{For scalar:}\quad & \hat{P}\varphi(x)\hat{P}^{-1} = +\varphi(\tilde{x}) \\ \text{For pseudoscalar:}\quad & \hat{P}\varphi(x)\hat{P}^{-1} = -\varphi(\tilde{x}) \end{aligned} \tag{2.3.1}$$

For the annihilation and creation operators a, a^+, b, b^+ defined in the following Fourier transformation

$$\varphi(x) = \int \frac{\mathrm{d}^3k}{(2\pi)^3\, 2k^0}\left(a(k)\mathrm{e}^{-\mathrm{i}k\cdot x} + b^+(k)\mathrm{e}^{\mathrm{i}k\cdot x}\right), \tag{2.3.2}$$

from Eqs. (2.3.1) and (2.3.2), we can derive the transformation of the annihilation and creation operators as

$$\begin{aligned} \text{For scalar:}\ & \hat{P}a(k)\hat{P}^{-1} = a(\tilde{k}), \quad \hat{P}a^+(k)\hat{P}^{-1} = a^+(\tilde{k}), \quad \tilde{k}^\mu = k_\mu \\ & \hat{P}b(k)\hat{P}^{-1} = b(\tilde{k}), \quad \hat{P}b^+(k)\hat{P}^{-1} = b^+(\tilde{k}) \\ \text{For pseudoscalar:}\ & \hat{P}a(k)\hat{P}^{-1} = -a(\tilde{k}), \quad \hat{P}a^+(k)\hat{P}^{-1} = -a^+(\tilde{k}), \quad \tilde{k}^\mu = k_\mu \\ & \hat{P}b(k)\hat{P}^{-1} = -b(\tilde{k}), \quad \hat{P}b^+(k)\hat{P}^{-1} = -b^+(\tilde{k}) \end{aligned}$$

---

[a] The $\theta \sim \tau$ puzzle, which will be further explained in subsequent chapter on weak interaction.

where $\tilde{k}^\mu = k_\mu$ means that the momentum of a particle $\vec{k}$ should be transformed to $-\vec{k}$ under space reflection.

2.3.2. *Parity transformation of fermion field*

To find the parity transformation of a fermion field $\psi(x)$, we assume the parity transformation of a fermion field is in the following form:

$$\psi'(x') = \Lambda\psi(x) \tag{2.3.3}$$

where Λ is the transformation matrix, and $x'^\mu = x_\mu$. Consider that Dirac equation should be invariant under parity transformation, then in parity transformed coordinate Dirac equation is

$$\left(\mathrm{i}\gamma^\mu \frac{\partial}{\partial x'^\mu} - m\right)\psi'(x') = 0 \tag{2.3.4}$$

Substitute Eq. (2.3.3) into Dirac equation

$$\mathrm{i}\gamma^\mu \frac{\partial x^\nu}{\partial x'^\mu}\frac{\partial}{\partial x^\nu}\Lambda\psi(x) - m\Lambda\psi(x) = 0 \tag{2.3.5}$$

Using $\partial x^\nu / \partial x'^\mu = g_{\mu\nu}$, the above equation is changed to be

$$\mathrm{i}\gamma_\nu \frac{\partial}{\partial x^\nu}\Lambda\psi(x) - m\Lambda\psi(x) = 0 \tag{2.3.6}$$

Equation (2.3.6) should be equivalent to Dirac equation, so Λ must be

$$\Lambda = c\gamma^0 \tag{2.3.7}$$

where c is a constant. To make the field $\psi(x)$ normalized, c must be a phase factor

$$c = \mathrm{e}^{\mathrm{i}\phi_\mathrm{P}} \tag{2.3.8}$$

Here ϕ_P is an unobservable phase angle. Therefore, the parity transformation of a fermion field is

$$\psi'(x') = \mathrm{e}^{\mathrm{i}\phi_\mathrm{P}}\gamma^0\psi(x) \tag{2.3.9}$$

Equation (2.3.9) can be also written as

$$\hat{P}\psi(x)\hat{P}^{-1} = \mathrm{e}^{\mathrm{i}\phi_\mathrm{P}}\gamma^0\psi(\tilde{x}) \tag{2.3.10}$$

Expand the field operator $\psi(x)$ with Fourier transformation

$$\psi(x)=\int\frac{\mathrm{d}^3k}{(2\pi)^3}\frac{m}{k^0}\sum_{\alpha=1,2}\left[a_\alpha(k)u^{(\alpha)}(k)\mathrm{e}^{-\mathrm{i}k\cdot x}+b_\alpha^+(k)v^{(\alpha)}(k)\mathrm{e}^{\mathrm{i}k\cdot x}\right] \tag{2.3.11}$$

where $a_\alpha(k)$ and $b_\alpha^+(k)$ are annihilation and creation operators of fermion and anti-fermion. One can prove that the parity transformations of the creation and annihilation operators are as follows:

$$\begin{aligned}\hat{P}a_\alpha(k)\hat{P}^{-1}&=\mathrm{e}^{\mathrm{i}\phi_\mathrm{P}}a_\alpha(\tilde{k}),\quad \hat{P}a_\alpha^+(k)\hat{P}^{-1}=\mathrm{e}^{-\mathrm{i}\phi_\mathrm{P}}a_\alpha^+(\tilde{k})\\ \hat{P}b_\alpha(k)\hat{P}^{-1}&=-\mathrm{e}^{-\mathrm{i}\phi_\mathrm{P}}b_\alpha(\tilde{k}),\quad \hat{P}b_\alpha^+(k)\hat{P}^{-1}=-\mathrm{e}^{\mathrm{i}\phi_\mathrm{P}}b_\alpha^+(\tilde{k})\end{aligned} \tag{2.3.12}$$

Note that parity transformation of a fermion field is different from its anti-fermion by a minus sign. Although there is an arbitrary phase for the parity transformation of a fermion field, the phase angle for the fermion is exactly opposite to that of the relevant anti-fermion. Therefore, the parity of a system composed of a fermion and its anti-fermion is independent of any arbitrary phase factor, i.e., the parity of the fermion and anti-fermion system is absolute, while the parity of a system of pure fermion or anti-fermions is arbitrary. The phase angle can be arbitrarily assigned.

In the following, for simplicity we use the convention for the arbitrary phase that $\phi_\mathrm{P}=0$. We shall discuss the parity transformation of a system composed of a fermion and its anti-fermion in the state ${}^{2S+1}L_J$. In the rest frame, the wave function of the system is

$$\left|{}^{2S+1}L_J\right\rangle=\sum_{m,\alpha,\beta}\int\mathrm{d}^3pY_{LM}(\vec{p})\chi_m(\alpha,\beta)C_{Mm}a_\alpha^+(\vec{p})b_\beta^+(-\vec{p})\left|0\right\rangle \tag{2.3.13}$$

where $Y_{LM}(\vec{p})$ is the spherical harmonics, which is the space wave function of the system. χ_m is the spin wave function, and C_{Mm} the Clebsch–Gordan coefficient for the spin–orbit coupling. A singlet state of two identical particles is antisymmetric under permutation of these two particles, while a triplet state is symmetric under permutation. Therefore, the symmetry property of a two-particle state is

$$\chi_m(\alpha,\beta)=(-1)^{S+1}\chi_m(\beta,\alpha) \tag{2.3.14}$$

For spherical harmonics, we have

$$Y_{LM}(-\vec{p})=(-1)^LY_{LM}(\vec{p}) \tag{2.3.15}$$

Acting the parity transformation operator in Hilbert space on the state $\left|{}^{2S+1}L_J\right\rangle$, we have

$$\begin{aligned}
\hat{P}\left|{}^{2S+1}L_J\right\rangle &= \sum_{m,\alpha,\beta}\int \mathrm{d}^3p Y_{LM}(\vec{p})\chi_m(\alpha,\beta)C_{Mm}\hat{P}a_\alpha^+(\vec{p})b_\beta^+(-\vec{p})\big|0\big\rangle \\
&= \sum_{m,\alpha,\beta}\int \mathrm{d}^3p Y_{LM}(\vec{p})\chi_m(\alpha,\beta)C_{Mm} \\
&\quad \times \hat{P}a_\alpha^+(\vec{p})\hat{P}^{-1}\hat{P}b_\beta^+(-\vec{p})\hat{P}^{-1}\hat{P}\big|0\big\rangle \\
&= -\sum_{m,\alpha,\beta}\int \mathrm{d}^3p Y_{LM}(\vec{p})\chi_m(\alpha,\beta)C_{Mm}a_\alpha^+(-\vec{p})b_\beta^+(\vec{p})\big|0\big\rangle \\
&\overset{\vec{p}\to-\vec{p}}{=} -\sum_{m,\alpha,\beta}\int \mathrm{d}^3p Y_{LM}(-\vec{p})\chi_m(\alpha,\beta)C_{Mm}a_\alpha^+(\vec{p})b_\beta^+(-\vec{p})\big|0\big\rangle \\
&= -(-1)^L\sum_{m,\alpha,\beta}\int \mathrm{d}^3p Y_{LM}(\vec{p})\chi_m(\alpha,\beta) \\
&\quad \times C_{Mm}a_\alpha^+(\vec{p})b_\beta^+(-\vec{p})\left|0\right\rangle \\
&= (-1)^{L+1}\left|{}^{2S+1}L_J\right\rangle \qquad (2.3.16)
\end{aligned}$$

From the second to the third step in the above equation, $\hat{P}\left|0\right\rangle = \left|0\right\rangle$ is used, which means that the vacuum state is invariant under parity transformation. Equation (2.3.16) implies that a state of a fermion and anti-fermion system with quantum number ${}^{2S+1}L_J$ is an eigenstate of parity transformation. For an S-wave state, $L = 0$, the parity of the system is negative.

2.3.3. *Parity transformation of vector field*

Considering the parity invariance of electromagnetic interaction $e\bar{\psi}\gamma_\mu\psi A^\mu$ in QED, parity transformation for the vector field $A_\mu(x)$ must be

$$\hat{P}A^\mu(x)\hat{P}^{-1} = A_\mu(\tilde{x}), \quad \tilde{x}^\mu = x_\mu \qquad (2.3.17)$$

which indicates that vector field does behave as a vector under parity transformation. Define annihilation and creation operators of photon a_λ, a_λ^+ through Fourier transformation of the vector field $A^\mu(x)$

$$A_\mu(x) = \int \frac{\mathrm{d}^3k}{(2\pi)^3\, 2k^0}\sum_{\lambda=0}^{3}\left[a_\lambda(k)\varepsilon_\mu^{(\lambda)}(k)\mathrm{e}^{-\mathrm{i}k\cdot x} + a_\lambda^+(k)\varepsilon_\mu^{(\lambda)*}(k)\mathrm{e}^{\mathrm{i}k\cdot x}\right] \qquad (2.3.18)$$

Using Eqs. (2.3.18) and (2.3.17), one can obtain parity transformation of creation and annihilation operators

$$\hat{P}a_\lambda(k)\hat{P}^{-1} = a_\lambda(\tilde{k}), \quad \hat{P}a_\lambda^+(k)\hat{P}^{-1} = a_\lambda^+(\tilde{k}) \tag{2.3.19}$$

simultaneously $\varepsilon_\mu^{(\lambda)}(\tilde{k}) = \varepsilon^{(\lambda)\mu}(k)$.

2.4. Charge Conjugation

Charge conjugation is the transformation between particle and antiparticle, which is denoted by the symbol C. The sign of charge of a particle shall be changed under charge conjugation. If a system is symmetric under charge conjugation, it is invariant under exchanges of particle and antiparticle.

2.4.1. *Charge conjugation of scalar field*

Charge conjugation of scalar field is defined as

$$\hat{C}\varphi(x)\hat{C}^{-1} = \varphi^+(x) \tag{2.4.1}$$

It is easily verified that Klein–Gordon equation is invariant under such a transformation. So it is a reasonable definition for the scalar field. The transformations of creation and annihilation operators of scalar field under charge conjugation are

$$\begin{aligned} &\hat{C}a(k)\hat{C}^{-1} = b(k), \quad \hat{C}a^+(k)\hat{C}^{-1} = b^+(k) \\ &\hat{C}b(k)\hat{C}^{-1} = a(k), \quad \hat{C}b^+(k)\hat{C}^{-1} = a^+(k) \end{aligned} \tag{2.4.2}$$

where $\hat{C}$ is a unitary operator. It satisfies

$$\hat{C}\,|0\rangle = |0\rangle \tag{2.4.3}$$

2.4.2. *Charge conjugation of fermion field $\psi(x)$*

Define charge conjugation of a fermion field $\psi(x)$ as

$$\hat{C}\psi(x)\hat{C}^{-1} = C\psi^C(x) \tag{2.4.4}$$

where C is the charge conjugate matrix, $\psi^C(x)$ the charge conjugated field of $\psi(x)$. To find C and $\psi^C(x)$, let us consider Dirac equation of fermion field $\psi(x)$ at first

$$(\mathrm{i}\gamma^\mu\partial_\mu - m)\,\psi(x) = 0 \tag{2.4.5}$$

Acting charge conjugation operator $\hat{C}$ and $\hat{C}^{-1}$ on both sides of the above equation, we obtain

$$(\mathrm{i}\gamma^\mu\partial_\mu - m)\,\hat{C}\psi(x)\hat{C}^{-1} = 0 \tag{2.4.6}$$

Substituting Eq. (2.4.4) into Eq. (2.4.6), we get

$$(\mathrm{i}\gamma^\mu\partial_\mu - m)\,C\psi^C(x) = 0 \tag{2.4.7}$$

Multiply $-C^{-1}$ from left to the above equation, then

$$\left(-\mathrm{i}C^{-1}\gamma^\mu C\partial_\mu + m\right)\psi^C(x) = 0 \tag{2.4.8}$$

In a further step, find Hermitian conjugate and transposition of Dirac equation

$$(\mathrm{i}\gamma^{\mu^\mathrm{T}}\partial_\mu + m)\bar{\psi}^\mathrm{T}(x) = 0 \tag{2.4.9}$$

Identify Eqs. (2.4.8) and (2.4.9), then one can obtain

$$\psi^C(x) = \bar{\psi}^\mathrm{T}(x) \tag{2.4.10}$$

and

$$C^{-1}\gamma^\mu C = -\gamma^{\mu^\mathrm{T}} \tag{2.4.11}$$

From Eq. (2.4.11) we can obtain

$$C = \mathrm{e}^{\mathrm{i}\theta_\mathrm{C}}\cdot \mathrm{i}\gamma^2\gamma^0 \tag{2.4.12}$$

where $\mathrm{e}^{\mathrm{i}\theta_\mathrm{C}}$ is the arbitrary phase factor. Considering both Eqs. (2.4.4) and (2.4.12), we can finally obtain the charge conjugate of the fermion field $\psi(x)$

$$\hat{C}\psi(x)\hat{C}^{-1} = \mathrm{e}^{\mathrm{i}\theta_\mathrm{C}}C\bar{\psi}^\mathrm{T}(x) \tag{2.4.13}$$

where the phase factor $\mathrm{e}^{\mathrm{i}\theta_\mathrm{C}}$ in the charge conjugate metrix C has been factored out, and the matrix C is redefined as

$$C = \mathrm{i}\gamma^2\gamma^0 \tag{2.4.14}$$

Using Eq. (2.4.13) and the Fourier transformation of $\psi(x)$, we can verify that the transformations of the creation and annihilation operators under charge conjugation are

$$\begin{aligned}\hat{C}a_\alpha(k)\hat{C}^{-1} &= \mathrm{e}^{\mathrm{i}\theta_{\mathrm{C}}}b_\alpha(k), \quad \hat{C}a_\alpha^+(k)\hat{C}^{-1} = \mathrm{e}^{-\mathrm{i}\theta_{\mathrm{C}}}b_\alpha^+(k)\\ \hat{C}b_\alpha(k)\hat{C}^{-1} &= \mathrm{e}^{-\mathrm{i}\theta_{\mathrm{C}}}a_\alpha(k), \quad \hat{C}b_\alpha^+(k)\hat{C}^{-1} = \mathrm{e}^{\mathrm{i}\theta_{\mathrm{C}}}a_\alpha^+(k)\end{aligned} \tag{2.4.15}$$

where $\alpha = \pm$ denotes the spin index of particles. Equation (2.4.15) means that charge conjugation does not change the direction of particle spin. Charge conjugation operator $\hat{C}$ is unitary operator, and vacuum does not charge under charge conjugation, $\hat{C}|0\rangle = |0\rangle$.

We usually take $\theta_c = 0$, so $\hat{C}\psi(x)\hat{C}^{-1} = C\bar{\psi}^{\mathrm{T}}(x)$. Similarly, we can derive $\hat{C}\bar{\psi}(x)\hat{C}^{-1} = -\psi^{\mathrm{T}}(x)C^{-1}$.

In the following we discuss the charge parity of a system composed of a fermion and its anti-fermion, which is in a state of ${}^{2S+1}L_J$. The quantum number of a system under charge conjugation is called charge parity.

Acting charge conjugation operator $\hat{C}$ on Eq. (2.3.13), we get

$$\hat{C}\left|{}^{2S+1}L_J\right\rangle = \sum_{m,\alpha,\beta}\int \mathrm{d}^3p Y_{LM}(\vec{p})\chi_m(\alpha,\beta)C_{Mm}\hat{C}a_\alpha^+(\vec{p})b_\beta^+(-\vec{p})\left|0\right\rangle \tag{2.4.16}$$

Since

$$\begin{aligned}\hat{C}a_\alpha^+(\vec{p})b_\beta^+(-\vec{p})\left|0\right\rangle &= \hat{C}a_\alpha^+(\vec{p})\hat{C}^{-1}\hat{C}b_\beta^+(-\vec{p})\hat{C}^{-1}\hat{C}\left|0\right\rangle\\ &= b_\alpha^+(\vec{p})a_\beta^+(-\vec{p})\left|0\right\rangle\end{aligned} \tag{2.4.17}$$

so Eq. (2.4.16) becomes

$$\begin{aligned}\hat{C}\left|{}^{2S+1}L_J\right\rangle &= \sum_{m,\alpha,\beta}\int \mathrm{d}^3p Y_{LM}(\vec{p})\chi_m(\alpha,\beta)C_{Mm}b_\alpha^+(\vec{p})a_\beta^+(-\vec{p})\left|0\right\rangle\\ &= -\sum_{m,\alpha,\beta}\int \mathrm{d}^3p Y_{LM}(\vec{p})\chi_m(\alpha,\beta)C_{Mm}a_\beta^+(-\vec{p})b_\alpha^+(\vec{p})\left|0\right\rangle\\ &\overset{\vec{p}\to-\vec{p}}{=} -\sum_{m,\alpha,\beta}\int \mathrm{d}^3p Y_{LM}(-\vec{p})\chi_m(\alpha,\beta)C_{Mm}a_\beta^+(\vec{p})b_\alpha^+(-\vec{p})\left|0\right\rangle\\ &\overset{\alpha\leftrightarrow\beta}{=} -\sum_{m,\alpha,\beta}\int \mathrm{d}^3p Y_{LM}(-\vec{p})\chi_m(\beta,\alpha)C_{Mm}a_\alpha^+(\vec{p})b_\beta^+(-\vec{p})\left|0\right\rangle\\ &= (-1)(-1)^L(-1)^{S+1}\sum_{m,\alpha,\beta}\int \mathrm{d}^3p Y_{LM}(\vec{p})\chi_m(\alpha,\beta)\\ &\quad \times C_{Mm}a_\alpha^+(\vec{p})b_\beta^+(-\vec{p})\left|0\right\rangle\\ &= (-1)^{L+S}\left|{}^{2S+1}L_J\right\rangle\end{aligned} \tag{2.4.18}$$

which means that the charge parity of fermion–anti-fermion system of $\left|{}^{2S+1}L_J\right\rangle$ is $(-1)^{L+S}$.

2.4.3. *Charge conjugation of vector field $A_\mu(x)$*

We discuss the case that $A_\mu(x)$ is electromagnetic field. In QED, the coupling of Fermion and electromagnetic field is $e\bar{\psi}\gamma_\mu\psi A^\mu$. The charge conjugation of fermion field has already been known. To keep the electromagnetic coupling in QED unchanged under charge conjugation, the charge conjugation of electromagnetic field must be

$$\hat{C}A_\mu(x)\hat{C}^{-1} = -A_\mu(x) \tag{2.4.19}$$

From Eq. (2.4.19) and the Fourier transformed expression of the electromagnetic field, we can obtain the charge conjugation of creation and annihilation operators of photon

$$\hat{C}a_\lambda(k)\hat{C}^{-1} = -a_\lambda(k), \quad \hat{C}a_\lambda^+(k)\hat{C}^{-1} = -a_\lambda^+(k) \tag{2.4.20}$$

From the above discussion, we can arrive at Furry theorem, which will be discussed in Section 2.5.

2.5. Application: Furry Theorem — A Particle of Spin 1 Cannot Decay to 2γ

2.5.1. *Furry theorem*

The wave function of a system composed of n photons is

$$\left|n\gamma\right\rangle \equiv \prod_{i=1}^{n} a_{\lambda_i}^+(k_i)\left|0\right\rangle \tag{2.5.1}$$

Perform C transformation for Eq. (2.5.1), then form Eq. (2.4.20) one can obtain

$$\hat{C}\left|n\gamma\right\rangle = (-1)^n\left|n\gamma\right\rangle \tag{2.5.2}$$

Consider the scattering of n photons into n' photons, $\left\langle n'\gamma\right|\hat{S}\left|n\gamma\right\rangle$, where $\hat{S}$ is scattering matrix. From C-invariance of QED, we have $\hat{S} = \hat{C}^+\hat{S}\hat{C}$, then

$$\begin{aligned}\left\langle n'\gamma\right|\hat{S}\left|n\gamma\right\rangle &= \left\langle n'\gamma\right|\hat{C}^+\hat{S}\hat{C}\left|n\gamma\right\rangle \\ &= (-1)^{n'+n}\left\langle n'\gamma\right|\hat{S}\left|n\gamma\right\rangle \end{aligned} \tag{2.5.3}$$

Therefore $n + n'$ must be even number. If $n + n'$ is odd, the scattering matrix element must be zero. So we can obtain Furry theorem:

$$\text{Photons of even numbers} \not\rightleftharpoons \text{photon of odd numbers}$$

Furry theorem can be also expressed as: the contribution of fermion loop attached with odd-number external photons in QED is zero [5].

2.5.2. *Particles of spin 1 cannot decay into two photons*

To show the application of the property of photons under parity transformation, let us consider a system composed of two photons. We can arrive at two meaningful results by studying the parity properties of this two-photon system [1]. We assume that these two photons emerge from a decay of an initial particle. At first let us clarify if two photons can be at a state with total angular momentum being 1.

If the total angular momentum of these two photons is 1, it must be a vector state. It behaves as a vector under coordinate rotation. Let us consider the wave function of this two-photon system in its center-of-mass frame. In general, its wave function can be written as

$$|2\gamma\rangle_i = \int \mathrm{d}^3p x_{ijk}(\vec{p})\varepsilon_j^{(\lambda_1)}(\vec{p})\varepsilon_k^{(\lambda_2)}(-\vec{p})a^+_{\lambda_1}(\vec{p})a^+_{\lambda_2}(-\vec{p})\,|0\rangle \tag{2.5.4}$$

where $i = 1,2,3$, is the vector index of this state. $x_{ijk}(\vec{p})$ is the space wave function in momentum space of this sate. It behaves as a tensor of rank-three under coordinate rotation. $x_{ijk}(\vec{p})$ can only depends on one variable $\vec{p}$. Its most general expression is

$$\begin{aligned} x_{ijk}(\vec{p}) = {} & A\varepsilon_{ijk} + Bp_i\delta_{jk} + Cp_j\delta_{ik} + Dp_k\delta_{ij} + Ep_i\varepsilon_{jkl}p_l \\ & + Fp_j\varepsilon_{kil}p_l + Gp_k\varepsilon_{ijl}p_l + Hp_ip_jp_k \end{aligned} \tag{2.5.5}$$

where the coefficients $A, \ldots, H$ are only scalar functions of momentum $|\vec{p}|$. Since $\vec{p}\cdot\varepsilon^{(\lambda)} = 0$, the terms of C, D, F, G and H do not contribute to the vector state. Therefore, Eq. (2.5.5) can be simplified to be

$$x_{ijk}(\vec{p}) = A\varepsilon_{ijk} + Bp_i\delta_{jk} + Ep_i\varepsilon_{jkl}p_l \tag{2.5.6}$$

Because photon is boson, its creation operators are commutable. So we can interchange the position of these two creation operators in Eq. (2.5.4) without changing the state of the two-photon system

$$\begin{aligned}|2\gamma\rangle_i &= \int \mathrm{d}^3p x_{ijk}(\vec{p})\varepsilon_j^{(\lambda_1)}(\vec{p})\varepsilon_k^{(\lambda_2)}(-\vec{p})a^+_{\lambda_2}(-\vec{p})a^+_{\lambda_1}(\vec{p})\,|0\rangle \\ &\overset{\vec{p}\to-\vec{p}}{=} \int \mathrm{d}^3p x_{ijk}(-\vec{p})\varepsilon_j^{(\lambda_1)}(-\vec{p})\varepsilon_k^{(\lambda_2)}(\vec{p})a^+_{\lambda_2}(\vec{p})a^+_{\lambda_1}(-\vec{p})\,|0\rangle \end{aligned} \tag{2.5.7}$$

Exchange the indices in the above equation $j \leftrightarrow k$, $\lambda_1 \leftrightarrow \lambda_2$

$$|2\gamma\rangle_i = \int \mathrm{d}^3p x_{ikj}(-\vec{p})\varepsilon_k^{(\lambda_2)}(-\vec{p})\varepsilon_j^{(\lambda_1)}(\vec{p})a^+_{\lambda_1}(\vec{p})a^+_{\lambda_2}(-\vec{p})\,|0\rangle \tag{2.5.8}$$

Adding Eqs. (2.5.4) and (2.5.8) and dividing the sum by 2, we can obtain a state exactly identical to the original state of the two-photon system

$$|2\gamma\rangle_i = \int \mathrm{d}^3p \frac{1}{2}\left[x_{ijk}(\vec{p}) + x_{ikj}(-\vec{p})\right]\varepsilon_j^{(\lambda_1)}(\vec{p})\varepsilon_k^{(\lambda_2)}(-\vec{p})a^+_{\lambda_1}(\vec{p})a^+_{\lambda_2}(-\vec{p})\,|0\rangle \tag{2.5.9}$$

From Eq. (2.5.6), we can see

$$x_{ikj}(-\vec{p}) = -x_{ijk}(\vec{p}) \tag{2.5.10}$$

So the vector state function of this photon system is zero, i.e.

$$|2\gamma\rangle_i = 0 \tag{2.5.11}$$

which means that it is impossible to construct a state with total angular momentum $J = 1$ for two-photon system. Therefore, we can draw a very significant conclusion:

(a) *Because of angular momentum conservation, a vector particle with spin $J = 1$ can not decay to two photons*

$$^{2S+1}L_{J=1} \nrightarrow 2\gamma \tag{2.5.12}$$

This selection rule is called Landau–Yang theorem [6, 7].

The reversed rule is also correct: if a particle can decay to two photons, its spin must not be 1. This can be used to determine the spin of a particle if it can decay to two photons.

Next let us discuss the state of two-photon system with spin $J = 0$.

A state with spin being 0 is a scalar state. The wave function of this state should be a scalar or pseudoscalar. So we need to construct a state without vector index. In general, we can write

$$|2\gamma\rangle = \int \mathrm{d}^3 p x_{ij}(\vec{p})\varepsilon_i^{(\lambda_1)}(\vec{p})\varepsilon_j^{(\lambda_2)}(-\vec{p})a_{\lambda_1}^{+}(\vec{p})a_{\lambda_2}^{+}(-\vec{p})\,|0\rangle \tag{2.5.13}$$

where $x_{ij}(\vec{p})$ is a tensor of rank-two. Similar to the method treating the above vector state, by considering the property under coordinate rotation, $x_{ij}(\vec{p})$ can be generally written as

$$x_{ij}(\vec{p}) = A\delta_{ij} + B\varepsilon_{ijk}p_k + Cp_ip_j \tag{2.5.14}$$

where the coefficients A, B and C are scalar functions of the momentum $|\vec{p}|$. Because of $\vec{p}\cdot\varepsilon^{\lambda} = 0$, the term C in Eq. (2.5.14) does not contribute, so we have

$$x_{ij}(\vec{p}) = A\delta_{ij} + B\varepsilon_{ijk}p_k \tag{2.5.15}$$

Substitute Eq. (2.5.15) into (2.5.13), we obtain

$$|2\gamma\rangle = \int \mathrm{d}^3 p \left[A\varepsilon^{(\lambda_1)}\cdot\varepsilon^{(\lambda_2)} + B\varepsilon^{(\lambda_1)}\times\varepsilon^{(\lambda_2)}\cdot\vec{p}\right] a_{\lambda_1}^{+}(\vec{p})a_{\lambda_2}^{+}(-\vec{p})\,|0\rangle \tag{2.5.16}$$

Acting the parity transformation operator on Eq. (2.5.16), we can see that the parity of the term A is 1, while the parity of term $B-1$, i.e. $\varepsilon^{(\lambda_1)}\cdot\varepsilon^{(\lambda_2)}$ is the scalar state of the two-photon system, and $\varepsilon^{(\lambda_1)}\times\varepsilon^{(\lambda_2)}\cdot\vec{p}$ the pseudoscalar state. Because parity is conserved in electromagnetic process, a scalar particle can only decays to two photons in a state with wave function $\varepsilon^{(\lambda_1)}\cdot\varepsilon^{(\lambda_2)}$, while a pseudoscalar particle can only decays to a state with wave function $\varepsilon^{(\lambda_1)}\times\varepsilon^{(\lambda_2)}\cdot\vec{p}$. Therefore we can draw the following conclusion

(b) *In the process of a scalar particle decaying into two photons, the polarizations of the two photons are mainly parallel. The decay amplitude is proportional to cos θ, with θ being the angle between the two photons' polarization vectors. In the process of a pseudoscalar particle decaying into two photons, the polarizations of the two photons are mainly perpendicular. The decay amplitude is proportional to sin θ.*

The above two conclusions (a) and (b) can be helpful to determine a particle's spin in experiment if it can decay into two photons [8].

2.6. Time Reversal

Time reversal is the transformation of changing the sign of time, i.e. $t \to -t$. The evolution sequence of physical events will reverse under the transformation of time reversal. An object flying from right to left will reverse direction, changing to be flying from left to right, under time reversal. An object dropping from the top of a building will rise to the top, etc. This is very similar to the operation of film backward-playing. Time reversal is discussed in physics, which does not mean that one can change the evolution direction of time in reality. It is only because physical laws in general will show significant properties under imaginary time reversal. One can draw significant conclusions from these properties. For example, Newton's equation is invariant under time reversal, which leads to the phenomenon that a process which happens in forward direction can also happens in backward direction actually. An object moving rightwards can certainly move leftwards. This is the consequence that mechanics law is invariant under time reversal. In addition, there are something that can happen at microlevel, but impossible at macrolevel. This is only because the probability in statistics is very tiny. For example, The time-reversal event of splashing water out of a bason is that water scattered everywhere returning back into the bason, which cannot happen in reality. This is not because the classical equation of motion violates the invariance of time reversal. In microlevel, water molecules can either move out of the bason, or move back into the bason. But, a huge amount of water molecules scattered everywhere moving back into the bason simultaneously, as a macroevent, cannot happen just because the statistics probability is too small.

Under time-reversal physical quantities which depend on time odd times will change sign, such as momentum $\vec{p} \to -\vec{p}$, angular momentum $\vec{L} = \vec{r} \times \vec{p} \to -\vec{L}$, etc. For spin, $\vec{\sigma}$, we assume that its behavior under time reversal is the same as that of angular momentum, so we have $\vec{\sigma} \to -\vec{\sigma}$. Quantities which do not depend on time or depend on time evenly remain unchanged under time reversal, such as, energy, electric charge, particle's helicity $\frac{\vec{\sigma}\cdot\vec{p}}{|\vec{p}|}$ etc., are all unchanged. Let us check the property of quantum mechanics under time reversal.

Define $\hat{T}$ to be time-reversal operator, and $\psi(\vec{x}, t)$ the wave function describing a particle's behavior in quantum mechanics. Then Schrödinger's equation is

$$\mathrm{i}\frac{\partial}{\partial t}\psi(\vec{x}, t) = H\psi(\vec{x}, t) \tag{2.6.1}$$

If the Hamiltonian for a particle is invariant under time reversal, then Schrödinger's equation should also be invariant under time reversal. Let us assume at first that time reversal is only the transformation of $t \to -t$.

Let us see what happens under this assumption. Under such a transformation, Schrödinger's equation is changed to be

$$-\mathrm{i}\frac{\partial}{\partial t}\psi(\vec{x},-t) = H\psi(\vec{x},-t) \tag{2.6.2}$$

where $\psi(\vec{x},-t)$ is the wave function after time reversal. Therefore, Schrödinger's equation is not invariant under such a naive transformation. Wigner suggested to alter the definition of time-reversal operation. The time-reversal operator $\hat{T}$ is modified to be not only changing the sign of time $t \rightarrow -t$, but also changing the wave function to be its complex conjugate [9]. Then under time reversal, Schrödinger's equation is changed to be

$$\begin{aligned}\mathrm{i}\frac{\partial}{\partial t}\psi^*(\vec{x},-t) &= H^*\psi^*(\vec{x},-t) \\ &= H\psi^*(\vec{x},-t) \qquad (2.6.3)\end{aligned}$$

If using $\hat{T}\psi(\vec{x},t) = \psi^*(\vec{x},-t)$ as the wave function after time reversal, and substituting it back into the above equation, we obtain

$$\mathrm{i}\frac{\partial}{\partial t}[\hat{T}\psi(\vec{x},t)] = H[\hat{T}\psi(\vec{x},t)] \tag{2.6.4}$$

After this modification, Schrödinger's equation will be invariant under time reversal. Although such a definition of time reversal looks weird, it does not violate any physical requirements in quantum mechanics. For example, conservation of probability under time reversal still holds

$$|\langle\psi\,|\psi\,\rangle|^2 = |\langle\hat{T}\psi|\hat{T}\psi\rangle|^2 \tag{2.6.5}$$

The time-reversal operator defined in such a way is not an unitary operator because of including the operation of complex conjugation. We can separate the complex conjugation in the definition of time-reversal operator as

$$\hat{T} \equiv \hat{U}\hat{K} \tag{2.6.6}$$

where $\hat{U}$ is unitary, which only takes the action of $t \rightarrow -t$ transformation, and $\hat{K}$ the operator performing complex conjugation. We shall prove that the time-reversal operator defined in Eq. (2.6.6) is not a unitary operator.

For any unitary operator $\hat{Y}$, when it acts on two states $|\alpha\rangle$ and $|\beta\rangle$, it satisfies

$$\langle \hat{Y}\alpha \,|\, \hat{Y}\beta\rangle = \langle \alpha \,|\, \hat{Y}^{+}\hat{Y} \,|\, \beta\rangle = \langle \alpha \,|\, \beta\rangle \tag{2.6.7}$$

We only needs to show that the operator $\hat{T}$ in Eq. (2.6.6) cannot be in according with the above equation. The proof is in the following

$$\begin{aligned}\langle \hat{T}\alpha \,|\, \hat{T}\beta\rangle &= \langle \hat{U}\hat{K}\alpha \,|\, \hat{U}\hat{K}\beta\rangle \\ &= \langle \hat{K}\alpha \,|\, \hat{U}^{+}\hat{U}\hat{K}\beta\rangle \\ &= \langle \hat{K}\alpha \,|\, \hat{K}\beta\rangle \\ &= [\langle \alpha \,|\, \beta\rangle]^{*}\end{aligned} \tag{2.6.8}$$

Equation (2.6.8) shows that $\hat{T}$ indeed does not satisfy Eq. (2.6.7). Therefore, it is not a unitary operator.

Hence $\hat{K}$ satisfies

$$\hat{K}\left(\lambda_1 \,|\alpha_1\rangle + \lambda_2 \,|\alpha_2\rangle\right) = \lambda_1^{*}\hat{K}\,|\alpha_1\rangle + \lambda_2^{*}\hat{K}\,|\alpha_2\rangle \tag{2.6.9}$$

Such kind of operators are not linear operators. They are antilinear operators. The operators as $\hat{U}\hat{K}$ are called as anti-unitary operators.

The time-reversal invariance of Schrödinger equation means that, if we find a solution of Schrödinger equation, we can obtain another time evolution function for a quantum state by acting time reversal on this solution, then the new quantum state is also solution of Schrödinger equation. What deserves to be noted is, to verify T invariance of a quantum system, it is not necessary to construct a quantum state of opposite time-evolution in experiment. It is impossible to do so either, because there are also evolution factors for phases in a quantum state except for evolutions of energy and momenta. Even if we can arrange the change of the directions of momentum and energy currents, we cannot precisely arrange the phases for the quantum state to keep it exactly to be the time-reversal state of the former. To verify T invariance of a quantum system, it is not by constructing a quantum state of opposite time-evolution in experiment, but by measuring any physical quantities and comparing the difference between time-reversal and un-reversal processes. For example, we can measure the probability for a neutral K meson $\bar{\mathrm{K}}^0$ transferring to its antiparticle K^0 at any time t for a neutral K meson system, and compare the difference of the transition probability for a reverse process. In such a process, the observable quantity is

$$\mathcal{A}_{\mathrm{T}} = \frac{\Gamma(\bar{\mathrm{K}}^0 \to \mathrm{K}^0, t) - \Gamma(\mathrm{K}^0 \to \bar{\mathrm{K}}^0, t)}{\Gamma(\bar{\mathrm{K}}^0 \to \mathrm{K}^0, t) + \Gamma(\mathrm{K}^0 \to \bar{\mathrm{K}}^0, t)}$$

where the function Γ is the transition width of the relevant particle transition process, which is proportional to the transition probability. For the measurement of T violation, the experimental result of CPLEAR Collaboration at CERN is $\mathcal{A}_{\mathrm{T}} = (6.6 \pm 1.3 \pm 1.0) \times 10^{-3}$ [10]. The details about neutral K meson system can be found in Chapter 7 of this book, or in Refs. [1,11].

The time-reversal operator $\hat{T}$ in Hilbert space can be obtained in quantum field theory, which is anti-unitary.

2.6.1. *Time reversal of Klein–Gordon field $\varphi(x)$*

Acting time-reversal operator on Klein–Gordon field $\varphi(x)$, one should have

$$\hat{T}\varphi(\vec{x}, t)\hat{T}^{-1} = \pm\varphi(\vec{x}, -t) \tag{2.6.10}$$

Complex conjugation operator $\hat{K}$ is included in $\hat{T}$, i.e.

$$\hat{T} = \hat{U}\hat{K} \tag{2.6.11}$$

here $\hat{U}$ is unitary operator. From Eqs. (2.6.10), (2.6.11) and the Fourier transformation of the field operator $\varphi(x)$, we can obtain that the annihilation and creation operators of Klein–Gordon field satisfy

$$\begin{aligned} \hat{U}a(k)\hat{U}^{-1} &= \pm a(\tilde{k}), \quad \hat{U}a^{+}(k)\hat{U}^{-1} = \pm a^{+}(\tilde{k}) \\ \hat{U}b(k)\hat{U}^{-1} &= \pm b(\tilde{k}), \quad \hat{U}b^{+}(k)\hat{U}^{-1} = \pm b^{+}(\tilde{k}) \end{aligned} \tag{2.6.12}$$

2.6.2. *Time reversal of fermion field*

Time-reversal operator of fermion field is also anti-unitary operator. It is also product of unitary operator and complex conjugation operator $\hat{T} = \hat{U}\hat{K}$. As an operator in Hilbert space, $\hat{T}$ satisfies

$$\hat{T}\psi(x)\hat{T}^{-1} = T\psi(-\tilde{x}) \tag{2.6.13}$$

where T is time-reversal matrix for the fermion field. Its explicit form can be obtained by requiring the invariance of the action of Dirac field under time reversal. Here the result is directly given below

$$T = \mathrm{i}\gamma^1\gamma^3 \tag{2.6.14}$$

In addition we can obtain the time reversal of annihilation and creation operators of fermion field

$$\begin{aligned}\hat{T}a_\alpha(k)\hat{T}^{-1} &= \alpha\mathrm{i}\mathrm{e}^{\mathrm{i}\theta_t}a_{-\alpha}(\tilde{k}), \quad \hat{T}a_\alpha^+(k)\hat{T}^{-1} = -\alpha\mathrm{i}\mathrm{e}^{-\mathrm{i}\theta_t}a_{-\alpha}^+(\tilde{k})\\ \hat{T}b_\alpha(k)\hat{T}^{-1} &= -\alpha\mathrm{i}\mathrm{e}^{-\mathrm{i}\theta_t}b_{-\alpha}(\tilde{k}), \quad \hat{T}b_\alpha^+(k)\hat{T}^{-1} = \alpha\mathrm{i}\mathrm{e}^{\mathrm{i}\theta_t}b_{-\alpha}^+(\tilde{k})\end{aligned} \tag{2.6.15}$$

where $\alpha = \pm$, is the helicity index of the particle, and θ_t an arbitrary phase angle.

2.6.3. *Time reversal of electromagnetic field $A_\mu(x)$*

From T invariance of electromagnetic interaction, one can obtain the transformation of electromagnetic field under time reversal

$$\hat{T}A_\mu(x)\hat{T}^{-1} = A^\mu(-\tilde{x}) \tag{2.6.16}$$

Proof. The coupling of electromagnetic field with fermion field is

$$\mathcal{L}_{\mathrm{em}} = \int \mathrm{d}^4x\mathrm{e}\bar{\psi}(x)\gamma_\mu\psi(x)A^\mu(x) \tag{2.6.17}$$

Perform T transformation for $\mathcal{L}_{\mathrm{em}}$

$$\hat{T}\mathcal{L}_{\mathrm{em}}\hat{T}^{-1} = \int \mathrm{d}^4x\mathrm{e}\psi^+(-\tilde{x})T^+\gamma^{0*}\gamma_\mu^*T\psi(-\tilde{x})\hat{T}A^\mu(x)\hat{T}^{-1} \tag{2.6.18}$$

Using $T^+\gamma_\mu^*T = \gamma^\mu$, Eq. (2.6.18) becomes

$$\hat{T}\mathcal{L}_{\mathrm{em}}\hat{T}^{-1} = \int \mathrm{d}^4x\mathrm{e}\bar{\psi}(-\tilde{x})\gamma^\mu\psi(-\tilde{x})\hat{T}A^\mu(x)\hat{T}^{-1} \tag{2.6.19}$$

From Eq. (2.6.19) we can see that, as long as $A_\mu(x)$ transforms as Eq. (2.6.16) under T transformation, we have

$$\mathcal{L}_{\mathrm{em}} = \hat{T}\mathcal{L}_{\mathrm{em}}\hat{T}^{-1} \tag{q.e.d}$$

With Eq. (2.6.16), and the Fourier transformation of $A_\mu(x)$

$$A_\mu(x) = \int \frac{\mathrm{d}^3k}{(2\pi)2k^0}\sum_\alpha (a_\alpha(k)\varepsilon_\mu^{(\alpha)}(k)\mathrm{e}^{-\mathrm{i}k\cdot x} + a_\alpha^+(k)\varepsilon_\mu^{(\alpha)^*}(k)\mathrm{e}^{\mathrm{i}k\cdot x}) \tag{2.6.20}$$

we can obtain T transformation of the annihilation and creation operators

$$\hat{T}a_\alpha(k)\hat{T}^{-1} = a_\alpha(\tilde{k}), \quad \hat{T}a_\alpha^+(k)\hat{T}^{-1} = a_\alpha^+(\tilde{k}) \tag{2.6.21}$$

□

In the following, we shall prove two theorems with time-reversal operator T discussed above [1].

(1) *If T invariance holds, then the $\hat{S}$ matrix for the evolution of a quantum state satisfies*

$$\hat{T}\hat{S}\hat{T}^{-1} = \hat{S}^+ \tag{2.6.22}$$

Proof. Let us recall the definition of $\hat{S}$ matrix at first. Let $|t_0\rangle$ and $|t\rangle$ be two quantum states at initial time t_0 and any time t, respectively. Then we can define the time-evolution operator through the relation of these two states:

$$|t\rangle = \hat{U}(t, t_0)\,|t_0\rangle \tag{2.6.23}$$

and $\hat{S}$ matrix is defined as

$$\hat{S} = \hat{U}(\infty, -\infty) \tag{2.6.24}$$

According to quantum mechanics, in interaction representation the time-evolution operator $\hat{U}(t, t_0)$ satisfies

$$\mathrm{i}\frac{\partial}{\partial t}\hat{U}(t, t_0) = \hat{H}_I(t)\hat{U}(t, t_0) \tag{2.6.25}$$

where $\hat{H}_I(t)$ is the Hamiltonian for the interacting part in the interaction representation. The initial condition is $\hat{U}(t_0, t_0) = 1$. Perform time-reversal transformation on both sides of Eq. (2.6.25),

$$-\mathrm{i}\frac{\partial}{\partial t}\hat{T}\hat{U}(t, t_0)\hat{T}^{-1} = \hat{T}\hat{H}_I(t)\hat{T}^{-1}\hat{T}\hat{U}(t, t_0)\hat{T}^{-1} \tag{2.6.26}$$

On the other hand, T invariance of the interaction means that the free and interacting parts of the Hamiltonian in Schrödinger representation satisfies

$$\hat{T}\hat{H}_0\hat{T}^{-1} = \hat{H}_0, \quad \hat{T}(\hat{H}_I)_S\hat{T}^{-1} = (\hat{H}_I)_S \tag{2.6.27}$$

where $\hat{H}_0$ and $(\hat{H}_I)_S$ denote free and interacting parts of the Hamiltonian in Schrödinger representation, respectively. In addition with the definition of $\hat{H}_I(t)$

$$\hat{H}_I(t) = \mathrm{e}^{\mathrm{i}\hat{H}_0 t}(\hat{H}_I)_S\mathrm{e}^{-\mathrm{i}\hat{H}_0 t} \tag{2.6.28}$$

one can obtain

$$\hat{T}\hat{H}_I(t)\hat{T}^{-1} = \hat{H}_I(-t) \tag{2.6.29}$$

Substituting Eq. (2.6.29) into (2.6.26), we have

$$\mathrm{i}\frac{\partial}{\partial(-t)}\hat{T}\hat{U}(t,t_0)\hat{T}^{-1} = \hat{H}_I(-t)\hat{T}\hat{U}(t,t_0)\hat{T}^{-1} \tag{2.6.30}$$

The above equation implies

$$\hat{T}\hat{U}(t,t_0)\hat{T}^{-1} = \hat{U}(-t,-t_0) \tag{2.6.31}$$

Therefore, from the definition of $\hat{S}$ matrix, we obtain

$$\begin{aligned}\hat{T}\hat{S}\hat{T}^{-1} &= \hat{U}(-\infty,\infty) = \hat{U}(\infty,-\infty)^{-1}\\ &= \hat{U}(\infty,-\infty)^{+}\\ &= \hat{S}^{+}\end{aligned}$$

□

(2) *Reciprocity*

For a scattering process

$$a + b + \cdots \rightarrow a' + b' + \cdots \tag{2.6.32}$$

let $|p_i,s_i\rangle$ and $|p'_j,s'_j\rangle$ denote the initial and final states, respectively, where p_i, s_i and p_j, s_j $(i = a,b,\ldots; j = a',b',\ldots)$ are momenta and spins of the initial and final states. If the interaction in the scattering process is T invariant, then according to theorem (1), the scattering $\hat{S}$ matrix satisfies

$$\hat{T}\hat{S}\hat{T}^{-1} = \hat{S}^{+} \tag{2.6.33}$$

Using Eq. (2.6.33), we can prove the reciprocity relation

$$\left|\langle p'_j,s'_j|\,\hat{S}\,|p_i,s_i\rangle\right| = \left|\langle -p_i,-s_i|\,\hat{S}\,|-p'_j,-s'_j\rangle\right| \tag{2.6.34}$$

Proof. The transformation of a particle state under time reversal is

$$\hat{T}\,|p_i,s_i\rangle = \mathrm{e}^{\mathrm{i}\theta}\,|-p_i,-s_i\rangle \tag{2.6.35}$$

$$\hat{T}\,|p'_j,s'_j\rangle = \mathrm{e}^{\mathrm{i}\theta'}\,|-p'_j,-s'_j\rangle \tag{2.6.36}$$

where $\mathrm{e}^{\mathrm{i}\theta}$ and $\mathrm{e}^{\mathrm{i}\theta'}$ are arbitrary phase factors. Insert the unit product $\hat{T}^{-1}\hat{T} = 1$ in the scattering amplitude,

$$\begin{aligned}\langle p'_j,s'_j|\,\hat{S}\,|p_i,s_i\rangle &= \langle p'_j,s'_j|\,\hat{T}^{-1}\hat{T}\hat{S}\hat{T}^{-1}\hat{T}\,|p_i,s_i\rangle\\ &= \langle p'_j,s'_j|\,\hat{T}^{+}\hat{T}\hat{S}\hat{T}^{-1}\hat{T}\,|p_i,s_i\rangle^{*}\end{aligned} \tag{2.6.37}$$

where we considered $\hat{T}$ is an anti-unitary operator in the second step, $\hat{T}^{-1} \neq \hat{T}^{+}$, and complex conjugation operation is included in $\hat{T}$. Then considering Eqs. (2.6.33), (2.6.35) and (2.6.36), we get

$$\begin{aligned}\left\langle p_j', s_j'\right| \hat{S} \left|p_i, s_i\right\rangle &= \mathrm{e}^{\mathrm{i}\theta'}\mathrm{e}^{-\mathrm{i}\theta}\left\langle -p_j', -s_j'\right| \hat{S}^{+} \left|-p_i, -s_i\right\rangle^{*} \\ &= \mathrm{e}^{\mathrm{i}\theta'}\mathrm{e}^{-\mathrm{i}\theta}\left\langle -p_i, -s_i\right| \hat{S} \left|-p_j', -s_j'\right\rangle \end{aligned} \tag{2.6.38}$$

So we can obtain the reciprocity relation, which is very meaningful in physics. It can relate the amplitudes of two inverse scattering processes with momenta and spins in opposite directions. Its applications will be discussed in subsequent chapters. □

2.7. *CPT* Theorem

In the preceding sections we discussed P, C and T transformations for scalar, spinor and vector fields. The interactions in nature do not always respect symmetries under P, C, T transformations exactly. Some products of any two kinds of these transformations are not exact symmetric transformation either. But if some conditions are satisfied, the joint transformation of P, C and T, no matter what the sequence is, is always a symmetric transformation. That is, the interaction is invariant under the joint transformation of P, C and T.

The operator for the joint transformation of C, P and T is defined as

$$\hat{I} \equiv \hat{C}\hat{P}\hat{T} \tag{2.7.1}$$

Then the CPT theorem can be stated as: if a theory is Lorentz invariant, and fields with spins of integer are quantized with commutation relation, while fields with spins of half-integer are quantized with anti-commutation relation, then the Lagrangian of this theory $\mathcal{L}(x)$ under the joint transformation of C, P and T transforms as

$$\hat{I}\mathcal{L}(x)\hat{I}^{-1} = \mathcal{L}^{+}(-x) = \mathcal{L}(-x) \tag{2.7.2}$$

Therefore, the action of this theory $S = \int \mathrm{d}^4x\mathcal{L}(x)$ is invariant. That is, the theory is invariant under the joint transformation of C, P and T.

The proof of this theorem is omitted here. The readers can refer to books of quantum field theory.

If the Hamiltonian of a system is invariant under the joint transformation of C, P and T, what kind of result can we obtain?

CPT invariance of the Hamiltonian $\hat{H}$ means: $\hat{I}\hat{H}\hat{I}^{-1} = \hat{H}$.

Table 2.2. The behavior of weak, electromagnetic and strong interaction under C, P, T transformations.

	P	C	T
Strong interaction	Conserved	Conserved	Conserved
Electromagnetic interaction	Conserved	Conserved	Conserved
Weak interaction	Violated	Violated	Violated

Let us have a look at the relation of the masses of a particle and its antiparticle.

For a particle state $|f\rangle$, under the joint transformation of C, P and T, the particle state $|f\rangle$ will be changed to be $\left|\bar{f}\right\rangle$,

$$\hat{I}\,|f\rangle = \mathrm{e}^{\mathrm{i}\theta}\left|\bar{f}\right\rangle \tag{2.7.3}$$

where $\mathrm{e}^{\mathrm{i}\theta}$ is an arbitrary phase factor. The mass of the particle equals to the expectation value of the Hamiltonian $\hat{H}$ between the state function of the particle in the rest frame

$$m_f = \langle f|\,\hat{H}\,|f\rangle \tag{2.7.4}$$

Inserting $\hat{I}^{-1}\hat{I} = 1$ in Eq. (2.7.4), we get

$$\begin{aligned} m_f &= \langle f|\,\hat{I}^{-1}\hat{I}\hat{H}\hat{I}^{-1}\hat{I}\,|f\rangle \\ &= \langle f|\,\hat{I}^{+}\hat{I}\hat{H}\hat{I}^{-1}\hat{I}\,|f\rangle^{*} \\ &= \left\langle\bar{f}\right|\hat{H}\left|\bar{f}\right\rangle^{*} = m_{\bar{f}}^{*} = m_{\bar{f}} \end{aligned} \tag{2.7.5}$$

That means, with the invariance of the Hamiltonian, one can derive that the masses of particle and antiparticle are equal.

Similarly, we can draw the conclusions that the mean lifetimes of particle and antiparticle are equal, and the electric charges of particle and antiparticle are the same by absolute value, but different by a sign, etc.

Before the end of this chapter, we list the behavior of weak, electromagnetic and strong interaction under C, P and T transformations (see Table 2.2).

References

1. Lee T D. *Introduction to Particle Physics and Field Theory*. Science press, 1980.
2. Noether E. *Gott. Nachr.*, 1918, 235–257.
3. Zhou B R. *Quantum Field Theory*. Higher Education Press, 2007.

4. Lee T D, Yang C N. *Phys. Rev.*, 1956, 104: 254.
5. Furry W H. *Phys. Rev.*, 1937, 51: 125.
6. Landau L D. *Dokl. Akad. Nawk. USSR*, 1948, 60: 207.
7. Yang C N. *Phys. Rev.*, 1950, 77: 242.
8. Plano R, *et al. Phys. Rev. Lett.*, 1959, 3: 525.
9. Wigner E P. *Group Theory and Its Application to the Quantum Mechanics of Atomic Spectra*. Academic Press, 1959.
10. Angelopoulos A, *et al.* (CPLEAR Collab.) *Phys. Lett.*, 1998, B43: 444.
11. Du D S. *CP Violation*. Peking University Press, 2012.

This page intentionally left blank

Chapter 3

The Classification and Properties of Particles: Lepton and Hadrons

3.1. Four Types of Interactions

Our understanding of interactions of matter is closely related to the knowledge on the structure of matter. Gravitation and electromagnetic interactions are two kinds of interactions that govern macroscopic motion. They are long-range forces and are also the first forms of interactions that human beings recognized. Knowledge of electricity and magnetism can date back to ancient times. In ancient China, magnetic stones have been recorded since the first century AD. At that time, Chinese had already discovered that the magnet had such characteristics as poles and later used it to invent the compass. The study of electrical and magnetic phenomena by means of modern science and technology originated in the 15th to 16th century. The introduction of the concept of electric and magnetic fields is a great advance in understanding of electric and magnetic forces. It changed the idea of force as acting at a distance. Maxwell studied electromagnetic theory, unified electricity and magnetism with his equations. Electricity and magnetism are unified interaction. They are two aspects of electromagnetic force. So the macroworld is governed by two interactions. Gravitational interaction governs the motion of celestial bodies, and exists between all matters with masses. Electromagnetic interaction governs the motion of charged objects, electromagnetic conversion, and electromagnetic wave emission and absorption etc.

With the deepening understanding of the microstructure of matter, two new forms of interactions were gradually discovered. The discovery of electrons at the end of the 19th century opened the door to the atomic world. The discovery of atomic nuclei has led to a new form of interaction — strong interaction comes into view. Nucleus is made up of protons and neutrons, which are not elementary particles. They are made up of more basic particles called quarks. Protons and neutrons are collectively called nucleons. Quarks are bound in the nucleons by strong interactions, which are short-range forces. Its acting distance is only between adjacent nucleons, that is, about $10^{-15}\,\mathrm{m}$. The typical duration of the strong interaction is very short, as little as $10^{-23}\,\mathrm{s}$. That is, if a particle is created or decays by strong interactions, the time it takes to produce or decay is usually on the order of $10^{-23}\,\mathrm{s}$. The typical interaction time is inversely proportional to the strength of the interaction. So nuclear force is a very strong interaction. This can be seen by comparing it with electromagnetic interaction. The typical time of electromagnetic interaction can be estimated from the electromagnetic decay of particles. For example, meson π^0 mainly decays to two photons, i.e., $\pi^0 \to \gamma\gamma$. This is electromagnetic decay process. The lifetime of π^0 is about $10^{-16}\,\mathrm{s}$, from which it can be estimated that the typical time for electromagnetic decay is about $10^{-16}\,\mathrm{s}$. About 10^7 times longer than the typical time of strong interaction. So strong interaction is much stronger than electromagnetic interaction.

Another type of interaction was found in β decay of nucleus and in decays of some other particles. The decay times of these decays are much longer. The half-life periods of nuclei due to β decay range from a few seconds to a few years. The mean lifetime of meson $\pi^{\pm}$ is about $2.6\times10^{-8}\,\mathrm{s}$, and K_{s}^0 is about $0.89 \times 10^{-10}\,\mathrm{s}$. These action times are 10^5–10^{20} times longer than the typical times for electromagnetic or strong interactions. So the interaction that governs these decays is a very weak interaction, which is called weak interaction. The force range of weak interaction is even shorter. We now know that weak interaction is transmitted by intermediate bosons, which include $\mathrm{W}^{\pm}$ and Z^0. The masses of $\mathrm{W}^{\pm}$ and Z^0 are as large as 80 GeV and 90 GeV, respectively. So the acting range of weak interaction is $R \sim \hbar/MC \approx 1/400\,\mathrm{fm}$ $(1\,\mathrm{fm} = 10^{-15}\,\mathrm{m})$. It is much shorter than the acting range of strong interaction.

There are four kinds of interactions in nature. They are gravitational interaction, electromagnetic interaction, weak interaction and strong interaction. The magnitude of their strengths can be compared by their coupling constants. The known fine structure constant α is the coupling constant for electromagnetic interaction. For strong interaction there is a similar coupling constant α_{s}. The coupling constant of weak interaction is the Fermi constant G_{F}. By comparing these coupling constants, we can see the relative sizes of these four interactions.

(1) Strong interaction: Dimensionless coupling constant

$$\alpha_{\mathrm{s}} = \frac{g_{\mathrm{s}}^2}{4\pi\hbar c} \approx 1 - 10$$

(2) Electromagnetic interaction: Dimensionless coupling constant

$$\alpha = \frac{e^2}{4\pi\varepsilon_0\hbar c} = \frac{1}{137}$$

(3) Weak interaction: The coupling constant is the Fermi constant

$$\frac{G_{\mathrm{F}}}{(\hbar c)^3} = 1.166 \times 10^{-5}\,\mathrm{GeV}^{-2},$$

which is however not a dimensionless coupling constant. In order to compare with the coupling constants of strong and electromagnetic interactions, we consider β decay of nucleons, and multiply Fermi constant with the squared typical scale of such decay process — the static energy of nucleon $(M_{\mathrm{P}}C^2)^2 \approx 1\,\mathrm{GeV}^2$. Then we get

$$\frac{G_{\mathrm{F}}}{(\hbar c)^3}(M_{\mathrm{P}}C^2)^2 \approx 1 \times 10^{-5},$$

where M_{P} is the mass of nucleon.

(4) Gravitational interaction: The magnitude of the gravitational force between two mass points with mass M and distance r is reflected by Newtonian constant. The force is $\frac{G_{\mathrm{N}}M^2}{r^2}$. Comparing with the static electric force between two point-charges with the electric charge e and distance r, one can know that the dimensional relation $[G_N M^2] \sim [\frac{e^2}{4\pi\varepsilon_0}]$. Therefore the combined quantity $\frac{G_{\mathrm{N}}M^2}{4\pi\hbar c}$ is also dimensionless. We can take M to be the mass of nucleon, $M = 1\,\mathrm{GeV}$, so as to see its relative size to weak, electromagnetic and strong interactions when the mass is at the scale of the nucleon mass. After calculating with $M = 1\mathrm{GeV}$, one can get

$$\frac{G_{\mathrm{N}}M^2}{4\pi\hbar c} \approx 5 \times 10^{-40}$$

Therefore, the comparison results of dimensionless coupling constants of strong interaction, electromagnetic interaction, weak interaction and

Table 3.1. Comparison of characteristics of four interactions.

	Strong	Electromagnetic	Weak	Gravitational
Source	Color	Electric charge	Weak charge	Mass
Constant	$\alpha_{\mathrm{s}} = \frac{g_{\mathrm{s}}^2}{4\pi\hbar c}$ $\approx 1 - 10$	$\alpha = \frac{e^2}{4\pi\varepsilon_0\hbar c}$ $\approx 1/137$	$\frac{G_{\mathrm{F}}(M_{\mathrm{p}}c^2)^2}{(\hbar c)^3}$ $\approx 1 \times 10^{-5}$	$\frac{G_{\mathrm{N}}M^2}{4\pi\hbar c}$ $\approx 5 \times 10^{-40}$
Mediator	Gluon (g)	Photon (γ)	Intermediate Bosons ($\mathrm{W}^{\pm}, \mathrm{Z}^0$)	—
Typical time	10^{-23} s	10^{-16} s	10^{-10} s	—
Acting range	1 fm	∞	1/400 fm	∞

gravitational interaction are

$$1 : \frac{1}{137} : 1 \times 10^{-5} : 5 \times 10^{-40}$$

So gravity is very weak compared to the other three kinds of interactions. The effect of gravitational interaction can be ignored when studying the interaction of microscopic particles.

Finally, we list some characteristics of these four interactions in Table 3.1.

3.2. Lepton and Lepton Number Conservation

3.2.1. *Electron, muon and neutrino*

Electron is the first elementary particle that has ever been discovered by human beings. By the 1930s, subatomic particles such as electrons, protons and neutrons had been discovered. Photons as mediator of electromagnetic interaction had also been discovered. Electrons, protons and neutrons are all fermions with spin of 1/2, which behave according to Fermi–Dirac statistics. Photons are vector particles with spin of 1, which are bosons and satisfy Bose–Einstein statistics. Electrons are extremely light compared with protons and neutrons. The mass of electron is $m_{\mathrm{e}} \approx 0.51\,\mathrm{MeV}$, and the masses of the protons and neutrons are $m_{\mathrm{p}} \approx 938.3\,\mathrm{MeV}$ and $m_{\mathrm{N}} \approx 939.6\,\mathrm{MeV}$, respectively. Electrons are called lepton, and protons and neutrons are called baryon. Electron has a negative charge of unit one, and it participates in electromagnetic interactions. From the fact that electrons can be generated in β decay of nucleus, we can know that electrons also participate in weak interaction.

In 1936, Anderson and Neddermeyer discovered muon in the study of cosmic rays with cloud chamber [1–4]. Its mass is $m_\mu \approx 105.7\,\text{MeV}$, and it has a negative unit charge. It was originally thought to be the particle that was predicted by Yukawa to transmit the nuclear force. But later it turned out experimentally that it was hardly involved in nuclear reactions [5], therefore it cannot participate in strong interactions, and cannot be transmitter of strong interactions. Subsequent experiments have shown that it has many properties similar to electrons. It also has a spin of 1/2, which is fermion. It also has a negative charge of unit one, and its antiparticle has a positive charge. It also participates in electromagnetic and weak interactions. So it was classified into the lepton family.

Electrons and muons are always associated with other two kinds of particles under weak action, which are neutrinos. Neutrinos are hypothesized to exist in nature by Pauli in 1930 to explain the continuous energy distribution of electrons in β decays of nucleus.

At that time, it was found that electrons generated in β decays of nucleus had a continuous energy spectrum. If β decay is a two-body reaction as

$$^A_ZX \rightarrow ^A_{Z+1}Y + \mathrm{e}^-$$

then according to the law of conservation of energy and momentum, electrons should have a certain value of energy. The energy of the electrons should not be continuously distributed. So the two-body decay process contradicts the experimental results. To solve this problem, Pauli hypothesized that another particle was also emitted in β decay except for electron. This particle has a charge of zero, a mass of almost zero, and is only involved in weak interaction. This particle is neutrino. Neutrinos rarely react with anything else. It can easily pass through something as large as earth, so it is very difficult to detect experimentally. That is, β decay is actually a three-body decay process. The basic process is

$$\mathrm{n} \rightarrow \mathrm{p} + \mathrm{e}^- + \bar{\nu}_\mathrm{e}$$

where $\bar{\nu}_\mathrm{e}$ an antineutrino. Neutrinos are just undetected because of their peculiar properties. So the electrons can have a continuous energy distribution.

In addition, in some typical nuclear β decay, for example,

$$\mathrm{O}^{14} \rightarrow \mathrm{N}^{14*} + \mathrm{e}^+ + \nu_\mathrm{e}$$

the spins of the initial and final nuclei O^{14} and N^{14*} are all 0. In such kind of decays, the total angular momentum taken away by positron and neutrino is 0. But in the decay process

$$\mathrm{He}^6 \to \mathrm{Li}^6 + \mathrm{e}^- + \bar{\nu}_\mathrm{e}$$

the spin of the initial nucleus He^6 is 0, while the spin of the final nucleus Li^6 is 1. The total angular momentum taken away by the electron and antineutrino in this decay is 1. The spin of the electron is 1/2. So in order to keep that the total angular momentum carried away by the electron (positron) and antineutrino (neutrino) can be only 0 or 1, the spin of neutrino or antineutrino must be 1/2. That is to say, the experiment of the nucleus β decay tells us that in order to preserve the conservation of angular momentum, the neutrino's spin should be 1/2.

In weak decays, the neutrino associated with muon is ν_μ, which is called μ-type neutrino. The μ-type neutrinos can be generated in the following processes

$$\pi^+ \to \mu^+ + \nu_\mu$$
$$\pi^- \to \nu^- + \overline{\nu}_\mu$$

Though individual neutrinos are extremely difficult to detect, they can be launched in great numbers from nuclear reactors. Reines and Cowan finally spotted a neutrino in a nuclear reactor in 1953 [6–8]. When nuclear fission occurs in a reactor, a lot of neutrons are released. Free neutrons will decay to produce antineutrino through the process $\mathrm{n} \to \mathrm{p} + \mathrm{e}^- + \bar{\nu}_\mathrm{e}$. Reines and Cowan placed scintillator detectors made of hydrogen-rich liquid near the reactor, through the inverse β process $\bar{\nu}_\mathrm{e} + \mathrm{p} \to \mathrm{n} + \mathrm{e}^+$ to detect the antineutrinos emitted from the reactor. Positrons quickly annihilate with electrons in water, and two photons with an energy of 0.51 MeV each are emitted. While neutrons move slowly in the water, finally they are absorbed by the cadmium to emit photons again. The cadmium has been dissolved in the water in advance. Two related flash events are detected in the experiment, and the rate of flash is directly related to the rate of fission reaction in the reactor. This confirms that the detected flash is caused by neutrinos emitted from the reactor. The experiment suggests that neutrinos have been detected directly outside their point of origin.

Another problem was also solved in 1962 [9]: Are neutrinos ν_e and ν_μ the same neutrino? Experimentally, high-energy proton beams are used to hit the target nucleus (Be) to produce a large number of secondary particles $\pi^\pm$ and $\mathrm{K}^\pm$ mesons. $\pi^\pm$ and $\mathrm{K}^\pm$ continue to decay to produce $\mu^\pm$ and neutrino and/or antineutrinos. Hadrons, such as $\pi^\pm$ and $\mathrm{K}^\pm$, and muons are absorbed by the shield wall. Only neutrinos can enter the detector.

The following reaction is detected experimentally

$$\bar{\nu}_\mu + \mathrm{p} \to \mathrm{n} + \mu^+$$

while there is no reaction

$$\bar{\nu}_\mu + \mathrm{p} \to \mathrm{n} + \mathrm{e}^+$$

being found.

In addition, only the following reaction happens

$$\bar{\nu}_\mathrm{e} + \mathrm{p} \to \mathrm{n} + \mathrm{e}^+$$

in inverse β decay using $\bar{\nu}_\mathrm{e}$ generated in reactors. That is, the neutrinos associated with muon and electron are not the same particles, i.e., ν_μ and ν_e are not the same.

3.2.2. τ *and its neutrino* ν_τ

The particle τ was discovered in electron–positron colliders in SLAC (America) and DESY (Germany) in 1975

$$\mathrm{e}^+ + \mathrm{e}^- \to \tau^+ + \tau^-$$

τ decays to electron or muon very quickly

$$\begin{aligned} \tau^+ &\to \mathrm{e}^+ + \nu_\mathrm{e} + \bar{\nu}_\tau \\ &\to \mu^+ + \nu_\mu + \bar{\nu}_\tau \end{aligned}$$

The mass and lifetime of τ are $m_\tau \approx 1.776\,\mathrm{GeV}$ and $\tau_\tau \approx 2.9 \times 10^{-13}\,\mathrm{s}$, respectively.

More experiments show that the properties of τ are very similar to muon and electron. It is charged with an electric charge of negative unit one, only involved in electromagnetic and weak interactions, and it does not participate in strong interaction, so it is also classified into leptons. Nowadays leptons are not classified by their masses. They are defined as fermions that do not participate in strong interactions.

The neutrino associated with lepton τ has been observed in experiment since 2001 [10].

Now, let us list the members of leptons as follows:

$$\begin{pmatrix} \nu_\mathrm{e} \\ \mathrm{e} \end{pmatrix} \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix} \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix} + \text{the relevant antiparticles} \tag{3.2.1}$$

Table 3.2. Lepton properties.

Lepton	m	Spin	Lifetime	Antiparticle
e^-	0.51 MeV	1/2	$> 4.6 \times 10^{26}$ yr	e^+
μ^-	105.6 MeV	1/2	2.2×10^{-6} s	μ^+
τ^-	1.776 GeV	1/2	$2.9{\times}10^{-13}$ s	τ^+
ν_e	<3 eV	1/2		$\bar{\nu}_e$
ν_μ	<0.19 MeV	1/2		$\bar{\nu}_\mu$
ν_τ	<18.2 MeV	1/2		$\bar{\nu}_\tau$

Theoretically, different types of leptons are called different flavors. This is just the name used to mark lepton types and has nothing to do with taste in our everyday lives. In formula (3.2.1), the pair of leptons in each bracket is called a generation. Now three generations of leptons have been found experimentally. Some properties of leptons are listed in Table 3.2.

3.2.3. *Helicity of neutrinos*

Neutrinos are fermions with spin 1/2, which are described by Dirac equation and the solutions are Dirac spinors. We know that there are four solutions to the Dirac equation, which correspond to the states of particle and antiparticles with spin being ±1/2, respectively. Consider the eigenstate of helicity $\frac{\vec{\sigma}\cdot\vec{p}}{|\vec{p}|}$, where $\vec{\sigma}$ is particle spin, and $\vec{p}$ its momentum. A state with eigenvalue of helicity +1 is defined as the right-handed state, and a state with helicity eigenvalue −1 defined as left-handed state. Therefore, the total four possible states for a fermion are: (1) the particle state with helicity being +1, (2) the particle state with helicity −1, (3) antiparticle state with helicity +1, (4) antiparticle state with helicity −1. In the case that the mass of the particle is not zero, these four solutions are independent. Each solution corresponds to a possible state of the particle.

The Dirac equation for a fermion with spin 1/2 is

$$(\gamma^\mu \hat{p}_\mu - m)\psi = 0 \tag{3.2.2}$$

where $\hat{p}_\mu = i\partial_\mu$ is the operator for four-momentum. Equation (3.2.2) can be further written as

$$\gamma^0 \hat{E}\psi = \vec{\gamma}\cdot\hat{\vec{p}}\psi + m\psi \tag{3.2.3}$$

Multiply $\gamma_5\gamma^0$ on both sides of Eq. (3.2.3) by the left, and use

$$\vec{\Sigma} = \gamma_5\gamma^0\vec{\gamma} = \begin{pmatrix} \vec{\sigma} & \\ & \vec{\sigma} \end{pmatrix} \tag{3.2.4}$$

then we can obtain

$$\gamma_5 \hat{E}\psi = \vec{\Sigma} \cdot \hat{\vec{p}}\psi + m\gamma_5\gamma_0\psi \tag{3.2.5}$$

where $\vec{\Sigma}$ is the spin operator of a four-component spinor. Again multiplying γ_5 on both sides of Eq. (3.2.5) by the left, we obtain another equation

$$\hat{E}\psi = \vec{\Sigma} \cdot \hat{\vec{p}}\gamma_5\psi + m\gamma_0\psi \tag{3.2.6}$$

Add and subtract Eqs. (3.2.5) and (3.2.6) to obtain the following two equations

$$\hat{E}\psi_{\mathrm{R}} = \vec{\Sigma} \cdot \hat{\vec{p}}\psi_{\mathrm{R}} + m\gamma_0\psi_{\mathrm{L}} \tag{3.2.7}$$

$$\hat{E}\psi_{\mathrm{L}} = -\vec{\Sigma} \cdot \hat{\vec{p}}\psi_{\mathrm{L}} + m\gamma_0\psi_{\mathrm{R}} \tag{3.2.8}$$

where ψ_{L} and ψ_{R} are called left-handed and right-handed fields, respectively. Their definitions are

$$\psi_{\mathrm{L}} = \frac{1-\gamma_5}{2}\psi, \quad \psi_{\mathrm{R}} = \frac{1+\gamma_5}{2}\psi \tag{3.2.9}$$

Because of the mass term, Eqs. (3.2.7) and (3.2.8) are interrelated. But when the mass of the particle is zero, the two equations are independent of each other

$$\hat{E}\psi_{\mathrm{R}} = \vec{\Sigma} \cdot \hat{\vec{p}}\psi_{\mathrm{R}} \tag{3.2.10}$$

$$\hat{E}\psi_{\mathrm{L}} = -\vec{\Sigma} \cdot \hat{\vec{p}}\psi_{\mathrm{L}} \tag{3.2.11}$$

From Eq. (3.2.10) for the right-handed field ψ_{R}, it can be seen that for the positive energy solution, the particle state has positive helicity. That is, the particle is dextral; For the negative energy solution, the particle state has a negative helicity. That is, the antiparticle is sinistral. While for the particle state described by the left-handed field ψ_{L}, the particle has an opposite helicity. That is, the particle is sinistral, while antiparticle dextral. So when the mass of the particle is zero, as the case of neutrinos, there are two possibilities for the solution of the Dirac equation. One possibility is that neutrino is sinistral, and antineutrino dextral. Another possibility is that neutrino is dextral and antineutrino sinistral. If only left-handed neutrino exists in nature, after parity transformation, because of the change of helicity, left-handed neutrino would become right-handed neutrino that does not exist in nature. This would destroy parity conservation. This is a theoretical explanation for parity non-conservation.

Experiments in 1958 confirmed that neutrinos were left-handed [11]. Neutrinos are therefore theoretically accepted as left-handed Dirac particles with a mass of zero and a spin of 1/2.[a]

3.2.4. *Lepton number conservation*

In all the reactions involving leptons, it is observed that, if attributing each lepton with a quantum number +1 or −1, such number is also conserved. This quantum number is called lepton number, and this law of conservation is called the law of conservation of lepton number. The lepton number of each lepton is usually defined as +1, and the lepton number of antilepton is −1. The lepton numbers for each lepton is shown in Tables 3.3 and 3.4.

Leptons of different types have different lepton numbers. For example, the lepton number of electron is $L_{\mathrm{e}} = 1$, and the lepton number of muon is $L_\mu = 1$, etc.

So far, all the experiments have shown that lepton numbers are conserved in all interactions, and each lepton number is conserved separately.

Table 3.3. List of lepton number for leptons L.

	e^-	μ^-	τ^-	ν_{e}	ν_μ	ν_τ
L_{e}	1	0	0	1	0	0
L_μ	0	1	0	0	1	0
L_τ	0	0	1	0	0	1

Table 3.4. List of lepton number for antileptons.

	e^+	μ^+	τ^+	$\bar{\nu}_{\mathrm{e}}$	$\bar{\nu}_\mu$	$\bar{\nu}_\tau$
L_{e}	−1	0	0	−1	0	0
L_μ	0	−1	0	0	−1	0
L_τ	0	0	−1	0	0	−1

---

[a] About the true nature of neutrinos, it has been a mystery for decades. Strictly speaking, the above property should be regarded as hypothesis. Neutrinos could also be another form, i.e., if you view the neutrino's antiparticle to be itself, the Dirac equation also has a solution. Such a theory was proposed by Italian physicist Majorana. Such particles are called Majorana particles. In Majorana's theory, neutrinos and antineutrinos are the same kind of particle and can have masses.

See that in the following reactions

$$
\begin{aligned}
&\mathrm{n} \to \mathrm{p} + \mathrm{e}^- + \bar{\nu}_\mathrm{e}\\
&\pi^+ \to \mu^+ + \nu_\mu\\
&\pi^- \to \mu^- + \bar{\nu}_\mu\\
&\mu^- \to \mathrm{e}^- + \bar{\nu}_\mathrm{e} + \nu_\mu\\
&\mu^+ \to \mathrm{e}^+ + \nu_\mathrm{e} + \bar{\nu}_\mu\\
&\tau^+ \to \mu^+ + \nu_\mu + \bar{\nu}_\tau
\end{aligned}
$$

etc., lepton numbers L_e, L_μ and L_τ conserve separately. According to lepton number conservation, we can determine which process is impossible. For example, the following processes: $\pi^- \to \mu^- + \nu_\mathrm{e}$, $\mu^- + \nu_\mu$, etc., are all impossible, because they violate the law of of lepton number conservation.

3.3. Hadrons: Conservation of Baryon Numbers

By 1932 it was known that atomic nuclei were made up of protons and neutrons. By binding protons and neutrons within the nucleus, there should be a very strong force — strong interaction, to counteract the Coulomb repulsions between protons. By studying the decay and scattering of atomic nuclei, it is found that, for particles that travel at almost the speed of light across the scale of nucleus which are only a few femtometers ($1\,\mathrm{fm} = 10^{-15}\,\mathrm{m}$), the action time is only $10^{-23}\,\mathrm{s}$, and in such a short time the scattering probability is very large, which indicates that the interaction must be very strong. It is also concluded that two characteristics must exist in strong interactions: (i) The force range is very short. The range of a nuclear force is only a few femtometers, and it only acts on a few nearby nucleons. (ii) The force is charge independent. The strong interaction between the nucleons has nothing to do with electric charge. There is no difference in the nuclear forces between proton–proton, proton–neutron and neutron–neutron.

With the study of the nuclear force, π meson, K meson and Λ strange baryon were discovered later. After the 1950s, particle physics entered a period of vigorous development. Accelerators that can artificially generate new particles have been built up, that led to the discovery of more new particles, such as a series of resonances, all of which participate in strong interactions. They are called hadrons. Hadrons are classified into baryons and mesons. Baryons include protons, neutrons, and strange baryon Λ, etc. Mesons include pion, K meson, and resonances ρ, ω, ϕ, etc.[b] In this section π meson, nucleon, strange particle and resonances will be introduced in turn. The quantum numbers including spin and parity of these particles and methods of partial wave analysis for resonances will be explained.

3.3.1. π *meson*

The π meson is a kind of particle predicted by Japanese physicist Hideki Yukawa in 1935 to solve the problem of nuclear force [12]. Considering that electromagnetic interactions are transmitted by electromagnetic fields, Yukawa believes that the nuclear force is transmitted through a field about nuclear force. There is also quanta for nuclear force field. Photons are massless, and the electromagnetic force acts over an infinite distance. Accordingly, it can lead to a short-range for the nuclear force by selecting an appropriate mass for the quantum of the nuclear force field. Yukawa's idea is as follows: considering the Klein–Gordon equation for scalar field

$$\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\psi = \nabla^2\psi - \frac{m^2c^2}{\hbar^2}\psi \tag{3.3.1}$$

If only considering the case for a static field, then Eq. (3.3.1) becomes

$$\left(\nabla^2 - \frac{m^2c^2}{\hbar^2}\right)\psi = 0 \tag{3.3.2}$$

Comparing Eq. (3.3.2) with the static potential equation for a point electric charge

$$\nabla^2\psi = 0$$

Yukawa interpreted Eq. (3.3.2) as a static nuclear force potential equation. Equation (3.3.2) can be expressed in spherical coordinates

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\psi(r)\right) = \frac{m^2c^2}{\hbar^2}\psi(r) \tag{3.3.3}$$

Solving the above equation, we can obtain

$$\psi(r) = \frac{g_0}{4\pi r}\mathrm{e}^{-r/R}, \quad R = \frac{\hbar}{mc} \tag{3.3.4}$$

---

[b] Some resonances, such as Δ^{++}, Δ^{+}, Δ^{0}, Δ^{-}, are baryons.

where g_0 is a constant, which can be explained as the charge for the nuclear field at the origin $r = 0$.

It can be seen from Eq. (3.3.4) that the potential drops sharply when $r > R$. Therefore, this result conforms to the requirement that the nuclear force is a short-range force. Its force range is $R = \hbar/mc$. Based on the magnitude of $10^{-15}\,\text{m}$ for the nuclear force known at that time, it can be obtained that the mass for the quantum is about $m \approx 100 - 200\,\text{MeV}$. Therefore, the mass of the transmitter for the nuclear force should be in the range $100 - 200\,\text{MeV}$. In 1947, the meson pion was observed in cosmic rays [13]. The mass of pion is about 140 MeV, which is consistent with Yukawa's theory. Subsequent studies have shown that there are strong interactions between pion and nucleus. So it can be sure that π meson is the particle that transmits the nuclear force.

There are three types of pions in total. They are π^+, π^- and π^0. That mass of the charged pion is $m_{\pi^\pm} \approx 139.6\,\text{MeV}$, and the mass of neutral pion is $m_{\pi^0} \approx 134.9\,\text{MeV}$.

The possible reactions for pion and nucleon scattering are

$$\begin{aligned}
&\pi^- + \mathrm{p} \to \mathrm{n} + \pi^0 \\
&\pi^+ + \mathrm{d} \to \mathrm{p} + \mathrm{p} \\
&\pi^+ + \mathrm{n} \to \pi^0 + \mathrm{p}
\end{aligned}$$

where d is deuteron, which is made up of a proton and a neutron. The reverse reaction for the above scattering can also happen.

The spin and parity of pion can be determined through reactions of pion and nucleon.

3.3.1.1. *The spin of pion*

The spin of the charged pion can be determined through the scattering process

$$\mathrm{p} + \mathrm{p} \to \pi^+ + \mathrm{d} \tag{3.3.5}$$

and its reverse process

$$\pi^+ + \mathrm{d} \to \mathrm{p} + \mathrm{p} \tag{3.3.6}$$

Without measuring the polarization state of the particle, i.e., averaging the polarization of the initial state and summing the polarization of the final state, calculate the cross sections of these two scattering processes (about cross section, see Sec. 3.4), then one can get

$$\sigma_{\mathrm{pp}\to\pi^+\mathrm{d}} \propto \frac{p_\pi^2}{(2J_\mathrm{p}+1)^2}\sum_{s_i s_f}|\langle\pi^+\mathrm{d}|\hat{S}|\mathrm{pp}\rangle|^2 \tag{3.3.7}$$

and

$$\sigma_{\pi^+\mathrm{d}\to\mathrm{pp}} \propto \frac{1}{2}\frac{p_p^2}{(2J_\pi+1)(2J_\mathrm{d}+1)}\sum_{s_i s_f}|\langle\mathrm{pp}|\hat{S}|\pi^+\mathrm{d}\rangle|^2 \tag{3.3.8}$$

where p_π and p_d are the momenta of pion and deuteron in the center-of-mass frame, respectively. J_π, J_d and J_p are spins of pion, deuteron and proton, respectively. The factor about angular momentum in the denominator is the result of averaging over the spin of the initial state. $\sum_{s_i s_f}$ denotes summation over the polarization states of the initial and final particles. We have known that all C, P and T conserve in strong interaction. Therefore, according to Sec. 2.6, the processes in Eqs. (3.3.5) and (3.3.6) are reciprocal, i.e.,

$$|\langle\pi^+\mathrm{d}|\hat{S}|\mathrm{pp}\rangle| = |\langle\mathrm{pp}|\hat{S}|\pi^+\mathrm{d}\rangle|$$

Then, from Eqs. (3.3.7) and (3.3.8), we can obtain

$$\frac{\sigma_{\mathrm{pp}\to\pi^+\mathrm{d}}}{\sigma_{\pi^+\mathrm{d}\to\mathrm{pp}}} = \frac{2\,(2J_\pi+1)\,(2J_\mathrm{d}+1)}{(2J_\mathrm{P}+1)^2}\frac{p_\pi^2}{p_\mathrm{p}^2} \tag{3.3.9}$$

where the spins of proton and deuteron have been known, $J_\mathrm{p} = 1/2$ and $J_\mathrm{d} = 1$. Therefore, comparing Eq. (3.3.9) with experimental data on the cross section, we can obtain the spin of pion. The spin of pion was obtained according to the experimental data in 1951–1953, $J_\pi = 0$ [14–17].

The spin of π^0 can be obtained by analyzing the decay process $\pi^0 \to \gamma\gamma$. The result is $J_\pi = 0$ (the readers can think about this problem by self). So both the spins of the charged and neutral pions are zero.

3.3.1.2. *The parity of pion*

The parity of the charged pion can be determined through the process of π^- with zero angular momentum being captured by deuteron

$$\pi^- + \mathrm{d} \to \mathrm{n} + \mathrm{n}$$

The idea is that, because parity is conserved under strong action, by analyzing the spin S and the quantum number of the orbital angular momentum L of each particle in the initial and final states, calculate the total parity of the initial and final states respectively, and then compare the two values to get the parity of π^- meson.

Let us analyze the parity of the initial state at first. The parity of the initial state should be $p_\mathrm{i} = p_\pi\ p_\mathrm{d}(-1)^L$, where $(-1)^L$ is the relative parity of π^- and d. In such process $L = 0$. d is composed of p and n, for which the relative angular momentum of p and n has been known to be even numbers $2n$, so $p_\mathrm{d} = p_\mathrm{p}p_\mathrm{n}(-1)^{2n} = p_\mathrm{p}p_\mathrm{n}$. Therefore, the total parity of the initial state is $p_\mathrm{i} = p_\pi p_\mathrm{p}p_\mathrm{n}(-1)^L = p_\pi p_\mathrm{p}p_\mathrm{n}$.

Secondly, let us consider the parity of the final state. The parity of the final state is $p_\mathrm{f} = p_\mathrm{n}p_\mathrm{n}(-1)^l$, where l is the relative orbital angular momentum between the two neutrons, which can be obtained by using the conservation of angular momentum between the initial and final states, and the asymmetric property of identical fermions. Note that neutron is fermion. The total angular momentum of the initial state J_i is obtained by coupling the total spin of π^- and d with the relative orbital angular momentum. The spins are $S_\pi = 0$, $S_\mathrm{d} = 1$, and the relative orbital angular momentum between π^- and d is 0, so $J_\mathrm{i} = 1$. The angular momentum of the final state is obtained by coupling the total spin of the two neutrons $S_{2\mathrm{n}}$ and the relative orbital angular momentum l. The possible values of $S_{2\mathrm{n}}$ are 0 and 1. So to get the total angular momentum of the final state to be 1, the possible values of l should be 0, 1, 2. The symmetric factor under exchange of two fermions is $(-1)^L(-1)^{S+1}$, so $L + S$ must be even numbers. Therefore, to obtain the total angular momentum of the final state to be 1, the only possible quantum number of the two neutrons is $l = 1$, $S_{2\mathrm{n}} = 1$. With this quantum numbers, we can get $p_\mathrm{f} = p_\mathrm{n}p_\mathrm{n}(-1)$. Compare p_i and p_f, we obtain

$$p_\pi p_\mathrm{p}p_\mathrm{n} = p_\mathrm{n}p_\mathrm{n}(-1)$$

Remove one p_n from both sides of the above equation, then

$$p_\pi p_\mathrm{p} = p_\mathrm{n}(-1)$$

So, once the parities of proton and neutron are known, the parity of π meson can be obtained according to the above equation. Because fermions do not have absolute parity, we need to assign convenient parity values for proton and neutron. As we will see in the next chapter, proton and neutron have almost the same behavior in strong interaction when ignoring electromagnetic effects and their tiny mass difference, so we can set the parity for proton and neutron to the same value, usually +1. Therefore, under such a parity convention, the parity of the charged pion is obtained to be −1.

The parity of π^0 can be measured through the decay process $\pi^0 \to \gamma\gamma$. The readers can think about the detail by themselves again. The measured value for π^0 through this decay mode is also -1 [18].

Therefore, the spins of these three pions are all 0, and parities all -1.

Under charge conjugation, $\hat{C}\left|\pi^+\right\rangle \to \left|\pi^-\right\rangle$, so $\pi^\pm$ is not the eigenstate of C transformation. $\hat{C}\left|\pi^0\right\rangle \to c\left|\pi^0\right\rangle$, so π^0 is the eigenstate of C transformation. The C parity of π^0 can be determined from the decay of $\pi^0 \to \gamma\gamma$. Photon is the quantum of vector field. From Chapter 2, we have known that the C parity of photon is -1, and C parity is multiplying quantum number, which is conserved under electromagnetic interaction. Therefore, it can be determined that the C parity of π^0 is $+1$.

3.3.2. *Nucleon, antinucleon and baryon number conservation*

Nucleons include proton p and neutron n. The antinucleons are $\bar{\mathrm{p}}$ and $\bar{\mathrm{n}}$. Both nucleon and antinucleon are fermions with spin 1/2. Although neutron is electrically neutral, its antiparticle is not itself. The antiproton was discovered in experiment in 1955 [19], and the antineutron was discovered in 1956 right after the discovery of antiproton [20].

A large amount of experimental results show that proton is basically stable, and its lifetime is more than 10^{32} years. We know that the mass of nucleon is about 1000 MeV, which is much large than that of mesons. Particles such as nucleons are collectively called baryons. The masses of baryons are much larger than meson. There is enough energy for a baryon to decay into mesons from the point of view of energy conservation, but such a decay has never actually been found. Some kind of conservation principle seems to forbid such kind of decays. To explain this phenomenon, Ernst Carl Gerlach Stueckelberg, Swiss mathematician and physicist, proposed the concept of baryon number and the conservation of baryon number.

Denote baryon number as B, and assign the baryon number of the baryon to be 1, antibaryon -1. For example, the baryon numbers of proton and neutron are $B = 1$, and the baryon numbers of antiproton and antineutron, $B = -1$. For the other particles, such as meson, lepton, and photon, all the baryon numbers of them are 0. Suppose that the baryon number is additive quantum number, then it is found that the total baryon number of particles in the initial and final states in any reaction remains unchanged. This is the law of conservation of baryon numbers. The conservation of baryon number is of universal significance. No signs of damage of this conservation law are found in any interactions.

3.3.3. *Strange particles*

3.3.3.1. *The discovery of strange particles*

In 1947, Rochester and Butler discovered some 'V'-type events when they used cloud chambers to study cosmic rays [21]. Since neutral particles cannot leave tracks in the cloud chamber, only a forked tracks left by two charged particles were observed, see Fig. 3.1. By measuring the deflection curvature and ionization density of the tracks of charged particles in the cloud chamber with magnetic field, it can be inferred that one of the two terminal particles is proton and the other is π^- meson. By measuring the energy and momentum of the terminal particles, the mass of the neutral particle can be determined. The mass is apparently larger than that of proton. This is a new particle in addition to proton, neutron, electron, photon and π meson known at that time. It was called Λ later. The particle reaction observed by Rochester and Butler is actually

$$\Lambda \to \mathrm{p} + \pi^-$$

The discovery of the new particle attracted wide attention. Later, another class of "V" type event was found. By analyzing, the terminal particles were determined to be π^+ and π^-. The mass of this new particle is about 1000 times of the mass of electron, which was called θ^0 particle at that time. Its lifetime is approximately 10^{-10} s. This is now known as the short-lived neutral K meson K_s^0.

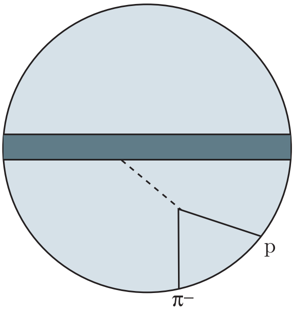


Fig. 3.1. Schematic diagram of "V" type event observed in cloud chamber.

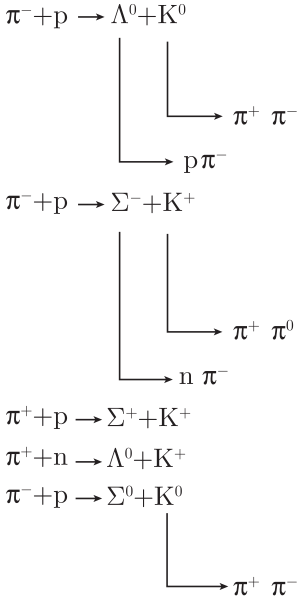
Soon afterwards, another new particle was found in the cosmic ray experiment, the mass of which is about 1000 times of that of electron. The particle is positively charged with a unit electric charge. It decays to $\pi^+\pi^0$ and/or $\mu^+\nu_\mu$. It is now called K^+ meson.

$$\mathrm{K}^+ \to \pi^+\pi^0, \quad \mathrm{K}^+ \to \mu^+\nu_\mu$$

In 1949, another decay mode of K^+ was found

$$\mathrm{K}^+ \to \pi^+\pi^+\pi^-$$

In 1953, 3 GeV proton synchrotron was built up in Brookhaven National Laboratory in the united states of America. Strange particles were then able to be generated artificially. Protons were accelerated to hit nuclear target, then beam of π meson was produced, the energy of which is as high as 1.5 GeV. The pion beam was used to hit other nuclear target to produce kinds of strange particles.



(3.3.10)

The strange particles can be divided into two categories according to their masses: the masses of one category are larger than that of pion, and smaller than nucleon, such as K^0, $\bar{\mathrm{K}}^0$, K^+, K^-, etc. The spins of this category are zero. They are bosons, and they are classified into mesons; The masses of the other category are larger than that of nucleons, such as Λ, Σ^+, Σ^0, Σ^-,

Ξ^-, Ξ^0, etc., which are called hyperon. The baryon number of hyperon is +1, and the baryon number of anti-hyperon is −1. The spins of these hyperons are 1/2. They are fermions. All these strange particles can be produced in strong interactions. They participate in strong interaction. So they are hadrons.

3.3.3.2. *Associated generation and strange number*

The properties of these particle are very strange. Their production cross section is at the order of millibarn. For example, the total cross section of pion of energy 1.5 GeV scattering at proton in the experiment of Brookhaven National Laboratory in 1953 is 34 ± 3 mb, but the cross section of the process producing strange particles such as $\pi^- + \mathrm{p} \to \Lambda^0 + \mathrm{K}^0$ is about 1 mb. The cross section of such magnitude is typically of scattering of strong interaction, i.e., the strange particles are produced in strong interaction. However, they decay much more slowly. By measuring the distance that Λ and K can move, the lifetime of them can be determined to be 2.6×10^{-10} s and 0.89×10^{-10} s, respectively, which are apparently the typical time that weak decay takes. Therefore, these strange particles are produced in strong interactions, but they only decay under weak interaction. That they can be generated in strong interaction implies they can participate in strong interaction. But, why do they only decay in weak interaction? It is very difficult to explain the huge difference between the creation and decay of these particles in theory. Actually before these particles can be produced in huge amount in accelerator, Pais has already proposed the idea of associated generation: these particles can only be generated in pair in strong interaction, while single particle can only decay in weak interaction [22]. Why can they only decay in weak interaction? It must be a sort of selection rule that forbids these particle to decay in strong interaction.

To explain the properties of strange particles, M. Gell-Mann, T. Nakano and K. Nishijima introduced the concept of strange number in 1953 [23,24]. They suppose that each particle has its strange number S. At first suppose that the strange numbers of ordinary particles are zero. For the strange particles, their strange numbers are determined according to the reactions they participate. Assign the strange number of one strange particle as $S = 1$, then the other particle generated associated with it must be $S = -1$. Assume strange number conserves in strong interaction, and can only be violated in weak interaction, then strange particles can decay to ordinary

Table 3.5. Strange numbers of some particles.

粒子	π	p	n	K^0	Λ	K^+	Σ^-	Σ^0
S	0	0	0	1	−1	1	−1	−1

particles only by weak interaction. In such a way, the associated generation of strange particles and the big difference between the generation and decay can be explained.

The strange numbers in Table 3.5 can be determined through the reactions in Eq. (3.3.10). We can assign strange number 1 for K^0 at first, then according to the possible process $\pi^- + p \to \Lambda + K^0$, the strange number of Λ can be determined to be −1. Using the process $\pi^+ + n \to \Lambda + K^+$ further, the strange number of K^+ can be known to be +1, etc.

Strange number is conserved in both strong and electromagnetic interactions, and be violated in weak interaction.

3.3.3.3. *The spin and parity of strange particles*

The spin and parity of strange particle can be determined by measuring the decay products of the strange particle and the reactions it participates, for example, $K^0 \to \pi^0\pi^0$, $K^+ \to \pi^+\pi^+\pi^-$, $\pi^- + p \to \Lambda^0 + K^0$, etc. Since angular momentum is strictly conserved, any process, no matter it proceeds under strong or weak interaction, can be used to determine the spin of strange particle. For any decay process, as long as the total spin of the final particles in the decay can be determined, the spin of the initial particle can be known according to angular momentum conservation. Parity is only conserved in strong and electromagnetic interactions, therefore the parity of a particle can only be determined through process under strong or electromagnetic interaction. Let us take $K^0 \to \pi^0\pi^0$ as an example to briefly show how to determine the spin of a particle.

According to quantum mechanics, the spacial wave function of a two-particle system is spherical harmonics. From the property of spherical harmonics, one can derive the symmetric factor of an identical two-particle system under exchange is $(-1)^l$, where l is the quantum number of the orbital angular momentum. π^0 is a boson of spin 0. For bosons the exchange symmetry requires that the symmetry factor $(-1)^l$ should be +1, so l must be even numbers. Analyzing the angular distribution of the decay product at the center-of-mass frame of K^0, no anisotropy is found. The spherical function can not be isotropic except for $l = 0$. Considering

Table 3.6. The spin and parity of strange particle.

	K^0	$\bar{K}^0$	K^+	K^-	Λ	Σ	Ξ
Spin	0	0	0	0	1/2	1/2	1/2
Parity	−1	−1	−1	−1	+1	+1	+1

the spin of π^0 is 0, we can determine the spin of K^0 is also 0. The spin and parity of some strange particles are listed in Table 3.6.

3.3.4. *Resonance*

From 1952 to 1954, a series of experiments of pion hitting liquid hydrogen target, performed by group led by Fermi in Chicago [25, 26], group of Lindenbaum and Luke Yuan at Brookhaven National Laboratory [27], and a group of Pennsylvania [28], showed that pion was more seriously absorbed by target particle than any other conditions when the energy of pion is at the region from 150 MeV to 200 MeV. This is very similar to the phenomenon of resonance of sound on an instrument. Playing specific notes at different pitches on an instrument, one finds that the notes that are played only within a certain range of frequencies have the highest volume. The frequency that causes the instrument to produce its loudest sound is the resonant frequency of the instrument. Resonance is a common phenomenon in many mechanical and electronic systems. Experiments of pion hitting liquid hydrogen target imply that there are also resonance phenomena in interacting particle system. The interacting system in a resonant state is called "resonance". Actually resonance is also a kind of particle. The only difference of them from ordinary particles is that they usually decay through strong interactions. So the lifetimes of them tend to be shorter, which is at the of order of typical time of strong interaction, 10^{-23} s. Therefore, resonance states are not fundamentally different from other particles.

In modern physics, particles are understood to be quanta of corresponding fields, or an eigenstate of an interacting system with a definite quantum number. The "resonance" usually refers to a state with a certain quantum number that appears with a greater probability near a certain energy. Therefore, there is no strict distinction between eigenstate, particle and resonance state.

With the rapid development of bubble chamber and electronic computer technology, the technique of particle detection has been greatly improved. Advances in technology have made it possible to find short-lived particles.

The tracks traveled by these short-lived particles are so short that they are technically impossible to measure directly. These resonance states are found by fitting the invariant mass spectra of the particles in the final state and analyzing the phase shift of the partial wave. The invariant mass spectrum refers to: at first, according to theory of special relativity, the four-momentum squared of a particle $p^2 = p_\mu p^\mu$ is Lorentz invariant quantity, the value of which is equal to the mass squared of the particle $m^2 = p^2$; Secondly, considering four-momentum conservation, $p = \sum_i p_i$, where p_i's are the four-momenta of particles in the final states. So one can obtain $p^2 = (\sum_i p_i)^2$. That is, by measuring the energies and momenta of particles in the final state, we can derive the mass of the initial particle. It is often found experimentally by fitting the energy and momentum of the final particles that an invariant mass usually forms a definite distribution — invariant mass spectrum. If a peak structure in the invariant mass spectrum appears at a certain value of mass, it means that there is a particle state near this mass.

In the 1950s and 1960s, a series of resonance states were found through the experiments of πN and KN scattering, such as baryon resonances Δ, Σ, Ω, etc., and meson resonances ρ, ω, ϕ and η, etc. [29–33]. In the scattering processes of πN and KN, one possibility is that πN or KN form a resonance state, and then this resonance state decays into other particles. Another possibility is to form a resonance state during the scattering of πN or KN, and in addition, there are several other particles produced simultaneously. In the former case, the resonance state can be found by the method of partial wave analysis. In the latter case, the existence of resonance state can be detected by fitting the invariant mass spectrum.

The method of partial wave analysis is introduced below.

Consider a beam of spin-0 particles bombarding a spin-0 target particle, and there is a potential around this target particle. Let us consider the case that this scattering is elastic scattering. Assume that the wave function of the beam particle is a plane wave that travels along Z-axis with momentum $\vec{k}$. The wave function of the incident wave is

$$\psi_i = \mathrm{e}^{\mathrm{i}kz} \tag{3.3.11}$$

The plane wave can be expanded into the form of spherical wave by using Legendre polynomial $p_l(\cos\theta)$, at large distance,

$$\psi_i(r,\theta) = \mathrm{e}^{\mathrm{i}kz} = \frac{\mathrm{i}}{2kr}\sum_{l=0}^{\infty}(2l+1)\left[(-1)^l\mathrm{e}^{-\mathrm{i}kr} - \mathrm{e}^{\mathrm{i}kr}\right]\cdot p_l(\cos\theta) \tag{3.3.12}$$

where θ is the polar angle between the radius vector $\vec{r}$ and Z-axis. The first term in Eq. (3.3.12) denotes the incoming wave with orbital quantum number l, and the second term the outgoing wave.

After the incident particle is scattered by the potential of the target particle, the phase and amplitude of the outgoing wave should be changed in principle. Let the phase change of the outgoing wave be $2\delta_l$ (the factor 2 is separated for the convenience of future discussion), and the amplitude changed to η_l ($1 \geq \eta_l \geq 0$ is called inelastic coefficient). Since the incoming wave is not affected by the potential, the wave function at distant place from the target particle should have the following form:

$$\psi(r,\theta) = \frac{\mathrm{i}}{2kr}\sum_{l=0}^{\infty}(2l+1)\left[(-1)^l \mathrm{e}^{-\mathrm{i}kr} - \eta_l \mathrm{e}^{2\mathrm{i}\delta_l}\mathrm{e}^{\mathrm{i}kr}\right]\cdot p_l(\cos\theta) \quad (3.3.13)$$

Subtracting Eq. (3.3.12) from Eq. (3.3.13), we can obtain the scattering wave function

$$\begin{aligned}\psi_{\mathrm{sc}}(r,\theta) &= \psi(r,\theta) - \psi_i(r,\theta)\\ &= \frac{\mathrm{e}^{\mathrm{i}kr}}{kr}\sum_l (2l+1)\frac{\eta_l \mathrm{e}^{2\mathrm{i}\delta_l}-1}{2\mathrm{i}}p_l(\cos\,\theta)\end{aligned} \quad (3.3.14)$$

Define the scattering amplitude as

$$f(k,\theta) = \frac{1}{k}\sum_l (2l+1)\frac{\eta_l \mathrm{e}^{2\mathrm{i}\delta_l}-1}{2\mathrm{i}}p_l(\cos\theta) \quad (3.3.15)$$

Eq. (3.3.14) can be written as

$$\psi_{\mathrm{sc}}(r,\theta) = f(k,\theta)\frac{\mathrm{e}^{\mathrm{i}kr}}{r} \quad (3.3.16)$$

The scattering amplitude $f(k,\,\theta)$ can be further written as

$$f(k,\theta) = \frac{1}{k}\sum_l (2l+1)f_l p_l(\cos\theta) \quad (3.3.17)$$

where f_l is called partial wave amplitude

$$f_l \equiv \frac{1}{2\mathrm{i}}(\eta_l \mathrm{e}^{2\mathrm{i}\delta_l} - 1) \quad (3.3.18)$$

According to the orthogonal relation of Legendre polynomials

$$\int p_l(\cos\theta)p_{l'}(\cos\theta)\mathrm{d}\Omega = \frac{4\pi\delta_{ll'}}{(2l+1)}, \quad (3.3.19)$$

where Ω is the solid angle, and the expression of the differential cross section of elastic scattering,

$$\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_{\mathrm{el}} = |f(k,\theta)|^2 , \tag{3.3.20}$$

the total cross section of elastic scattering is obtained by integrating over solid angle

$$\sigma_{\mathrm{el}} = 4\pi\bar{\lambda}^2 \sum_l (2l+1) \left|\frac{\eta_l \mathrm{e}^{2\mathrm{i}\delta_l} - 1}{2\mathrm{i}}\right|^2 \tag{3.3.21}$$

Here $\lambda = 1/k$. When $\eta_l = 1$, i.e., for the pure elastic scattering,

$$\sigma_{\mathrm{el}} = 4\pi\lambda^2 \sum_l (2l+1) \sin^2 \delta_l \tag{3.3.22}$$

The cross section of inelastic scattering can be obtained from conservation of probability. Define the first term of Eq. (3.3.14) as amplitude of elastic outgoing wave,

$$\psi_{\mathrm{eo}} = \frac{\mathrm{e}^{\mathrm{i}kr}}{kr} \sum_l (2l+1) \frac{\eta_l \mathrm{e}^{2\mathrm{i}\delta_l}}{2\mathrm{i}} p_l(\cos\theta) \tag{3.3.23}$$

The second term is the amplitude of the outgoing wave without scattering potential

$$\psi_{\mathrm{to}} = \frac{\mathrm{e}^{\mathrm{i}kr}}{kr} \sum_l (2l+1) \frac{1}{2\mathrm{i}} p_l(\cos\theta) \tag{3.3.24}$$

Then the inelastic scattering cross section is

$$\begin{aligned} \sigma_{\mathrm{n}} &= \int (|\psi_{\mathrm{to}}|^2 - |\psi_{\mathrm{eo}}|^2) r^2 \mathrm{d}\Omega \\ &= \pi\lambda^2 \sum_l (2l+1)(1 - |\eta_l|^2) \end{aligned} \tag{3.3.25}$$

The total cross section is

$$\sigma_{\mathrm{t}} = \sigma_{\mathrm{el}} + \sigma_{\mathrm{n}} = 2\pi\lambda^2 \sum_l (2l+1)(1 - \eta_l \cos 2\delta_l) \tag{3.3.26}$$

From Eq. (3.3.17) and $p_l(1) = 1$, we can obtain

$$\mathrm{Im} f(k,0) = \frac{1}{2k} \sum_l (2l+1)\,(1 - \eta_l \cos 2\delta_l) \tag{3.3.27}$$

Comparing Eqs. (3.3.26) and (3.3.27), the following relation can be obtained

$$\mathrm{Im} f(k,0) = \frac{k}{4\pi}\sigma_{\mathrm{t}} \tag{3.3.28}$$

This is called optical theorem. That is, the imaginary part of the forward scattering amplitude is proportional to the total scattering cross section.

It can be known from Eq. (3.3.22) that, when $\delta_l = \pi/2$, the maximum value is obtained for the cross section of the elastic scattering for l partial wave

$$\left(\sigma_{\mathrm{el}}^{l}\right)_{\max} = 4\pi\lambda\!\!\!^{-2}(2l+1) \tag{3.3.29}$$

From Eq. (3.3.25), it is known that the maximum value is obtained for the cross section of the inelastic scattering for l partial wave as $\eta_l = 0$,

$$\left(\sigma_{\mathrm{n}}^{l}\right)_{\max} = \pi\lambda\!\!\!^{-2}(2l+1) \tag{3.3.30}$$

Equations (3.3.29) and (3.3.30) are called unitary limit for partial waves.

The partial wave amplitude with orbital angular quantum number l in Eq. (3.3.18) can be also written as

$$f_l \equiv \frac{\mathrm{i}}{2} - \frac{\mathrm{i}\eta_l}{2}\mathrm{e}^{2\mathrm{i}\delta_l} \tag{3.3.31}$$

It can be known from Eq. (3.3.31) that f_l is a complex number, which can be represented in the complex plane. For $\eta_l = 1$, f_l varies with the phase shift δ_l, which becomes a unit circle on the complex plane. The center of the unit circle is at point (0, i/2). For $\eta_l < 1$, f_l forms a curve within the unit circle. Such diagram is called Argand diagram, see Fig. 3.2.

Partial wave phase-shift δ_l is generally dependent on the center-of-mass energy of the system E. We can explicitly write this dependence as $\delta_l(E)$. The maximum value for the partial wave amplitude is obtained as

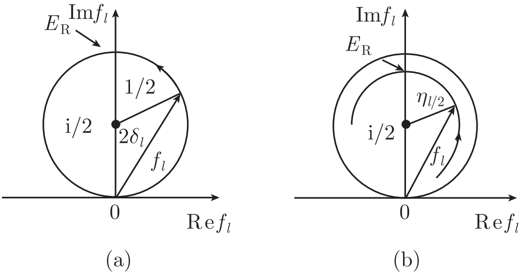


Fig. 3.2. Argand diagram: (a) for $\eta_\ell = 1$; (b) for $\eta_\ell < 1$.

$\delta_l(E) = \pi/2$, where the system forms a resonance state, and the center-of-mass energy corresponds to resonant energy $E = E_{\mathrm{R}}$ at this time. For $\eta_l = 1$, expand f_l around the resonant energy E_{R} in Taylor expansion. To do this, we rewrite f_l as

$$\begin{aligned} f_l &= \frac{\mathrm{i}}{2} - \frac{\mathrm{i}}{2}\mathrm{e}^{2\mathrm{i}\delta_l(E)} = \mathrm{e}^{\mathrm{i}\delta_l(E)}\sin\delta_l(E) \\ &= \frac{1}{\cot\delta_l(E) - \mathrm{i}} \end{aligned} \tag{3.3.32}$$

Expand $\cot\delta_l(E)$ around E_{R} at first

$$\cot\delta_l(E) = \cot\delta_l(E_R) + (E - E_R)\,\frac{\mathrm{d}}{\mathrm{d}E}\cot\delta_l(E)\big|_{E=E_{\mathrm{R}}} + \cdots \tag{3.3.33}$$

Since $\delta_l(E_{\mathrm{R}}) = \pi/2$, so cot $\delta_l(E_{\mathrm{R}}) = 0$. Define

$$\frac{\mathrm{d}}{\mathrm{d}E}\cot\delta_l(E)\,|_{E=E_{\mathrm{R}}} = \frac{2}{\Gamma} \tag{3.3.34}$$

Then we get

$$\cot\delta_l(E) \approx (E - E_{\mathrm{R}})\,\frac{2}{\Gamma} \tag{3.3.35}$$

Substituting Eq. (3.3.35) into Eq. (3.3.32), and ignoring higher-order terms, we get

$$f_l = \frac{\Gamma/2}{E - E_R - \mathrm{i}\Gamma/2} \tag{3.3.36}$$

From Eqs. (3.3.36) and (3.3.21), we can obtain the elastic scattering cross section for the lth partial wave

$$\sigma_{\mathrm{el}}^l = 4\pi\lambda\!\!\!^-{}^2(2l+1)\frac{\Gamma^2/4}{(E - E_{\mathrm{R}})^2 + \Gamma^2/4} \tag{3.3.37}$$

This is called Breit–Wigner formula. There is a resonance peak at $E = E_{\mathrm{R}}$ for the elastic cross section. The width at half height of the peak is Γ.

The method of partial wave analysis introduced above provides us with a method to find the short-lived resonance states. In the experiment of using $\pi^{\pm}$ meson hitting liquid hydrogen, measuring and depicting the curve of the πN scattering cross section varying with the center-of-mass energy of the system, apparent peak can be found around 1230 MeV, which is the clear evidence for the existence of a resonance state. This is the first resonance state Δ(1232). Δ(1232) is a P wave resonance of isospin $I = 3/2$ and spin $J = 3/2$. Isospin will be introduced in Chapter 4. In the literature, the symbol $l_{2I,2J}$ or $(2I,\ 2J)$ is usually used to denote a resonance, where l stands for the lth partial wave. For example, $\Delta(1232)$ can be denoted by P_{33} or (3,3) resonance. Four different electric states for $\Delta(1232)$ have been found in experiment, which are denoted as Δ^{++}, Δ^{+}, Δ^{0}, Δ^{-}, respectively. As for how to experimentally determine the quantum numbers such as spin and parity for resonance states, the readers can refer to some other relevant literatures or books.

In addition, in πp scattering experiment, by using the following processes

$$\begin{tikzcd}[column sep=small] \pi^+ + p \arrow[r] & \rho^+ + p \arrow[d] \\ & \pi^+ + \pi^0 \end{tikzcd}$$

$$\begin{tikzcd}[column sep=small] \pi^+ + p \arrow[r] & \rho^0 + p + \pi^+ \arrow[d] \\ & \pi^+ + \pi^- \end{tikzcd}$$

the resonance ρ has been found. There are three electric states for ρ resonance: ρ^+, ρ^0 and ρ^-. By analyzing the invariant mass spectrum of $\pi\pi$, the mass of ρ meson is determined to be 770 MeV [33]. The ρ meson can also be experimentally studied using another processes

$$\begin{tikzcd}[column sep=small] \pi^- + p \arrow[r] & \rho^0 + n \arrow[d] \\ & \pi^+ + \pi^- \end{tikzcd}$$

$$\begin{tikzcd}[column sep=small] \pi^- + p \arrow[r] & \rho^- + p \arrow[d] \\ & \pi^- + \pi^0 \end{tikzcd}$$

3.4. Scattering Cross Section, Particle Lifetime and Decay Width

3.4.1. *Scattering cross section*

Consider the process of a beam of particles 1 bombarding target particle 2

$$1 + 2 \to 1' + 2' + \cdots + n'$$

We treat this problem in the rest frame of the target particle at first, then transfer to any coordinate frame.

Assume that the particle number density of beam 1 is ρ_1, beam velocity v_1, and the target particles locate in a thin layer of cross section area S_2 and thickness dx. The density of the target particle is ρ_2.

If the effective scattering area of a single target particle is $\mathrm{d}\sigma$, then the total effective scattering area in the whole thin layer with cross section area S_2 and thickness $\mathrm{d}x$ should be $\mathrm{d}\sigma\rho_2 S_2\mathrm{d}x$. The probability of the beam particle moving into the volume $S_2\mathrm{d}x$ at time interval $\mathrm{d}t$ and being effectively scattered by the target particles is proportional to the product of particle numbers in the beam and that in the target at such time interval

$$\mathrm{d}p = c\rho_1 v_1 \mathrm{d}t \cdot \mathrm{d}\sigma\rho_2 S_2 \mathrm{d}x \tag{3.4.1}$$

where c is an proportional constant. If absorbing c into $\mathrm{d}\sigma$, which corresponds to redefining $\mathrm{d}\sigma$, then

$$\mathrm{d}\sigma = \frac{\mathrm{d}p}{\rho_1\rho_2 v_1 S_2 \mathrm{d}x\mathrm{d}t} \tag{3.4.2}$$

Define the scattering probability per unit space–time volume as

$$\mathrm{d}W = \mathrm{d}p/(S_2\mathrm{d}x\mathrm{d}t) \tag{3.4.3}$$

then Eq. (3.4.2) becomes

$$\mathrm{d}\sigma = \frac{\mathrm{d}W}{\rho_1\rho_2 v_1} \tag{3.4.4}$$

If transfer to the coordinate frame where the two particle collide, i.e., particle 2 also moves with a definite velocity v_2, and in addition v_1 and v_2 being antiparallel, then v_1 in Eq. (3.4.4) should be substituted by the relative velocity $v_{12} = |\bar{v}_1 - \bar{v}_2|$. So

$$\mathrm{d}\sigma = \frac{\mathrm{d}W}{\rho_1\rho_2 v_{12}} \tag{3.4.5}$$

where $\mathrm{d}\sigma$ in the left-handside of Eq. (3.4.5) is called differential cross section. The scattering probability in unit space–time volume $\mathrm{d}W$ can be calculated according to quantum field theory.

Since particles in a collision mostly pass through each other without scattering, therefore, the S matrix of the scattering process can be decomposed as

$$\hat{S} = \hat{I} + \mathrm{i}\hat{R}$$

where $\hat{I}$ is the operator of unit matrix describing the case without scattering, while $\hat{R}$ includes all the effect of the interaction during the scattering.

The S-matrix for the scattering can be written as

$$\begin{aligned}
&\langle 1', 2', \ldots, n'|\, \hat{S}\, |1, 2\rangle \\
&\quad = \langle 1', 2', \ldots, n'|\, \hat{I}\, |1, 2\rangle \\
&\qquad + \mathrm{i}(2\pi)^4\delta^4\left(\sum_{i=1}^{n} p_i' - p_1 - p_2\right)\langle 1', 2', \ldots, n'|\, \hat{T}\, |1, 2\rangle
\end{aligned} \tag{3.4.6}$$

where the second term in the above equation is the matrix element describing the scattering process of $1 + 2 \to 1' + 2' + \cdots + n'$, and the δ function is the four-momentum conservation factor. To calculate the scattering probability in unit space–time volume $\mathrm{d}W$, the second term of Eq. (3.4.6) should be multiplied with the quantum state number within momentum-space $\mathrm{d}^3p_1'\mathrm{d}^3p_2' \cdots \mathrm{d}^3p_n'$, which is also called phase space.

Take the convention for the normalization of a physical state as

$$\begin{array}{ll}
\text{For bosons:} & \langle p'|p\rangle = (2\pi)^3\, 2p^0\delta^3(\vec{p'} - p) \\
\text{For massive fermions:} & \langle p', \alpha|p, \alpha\rangle = (2\pi)^3\dfrac{p^0}{m}\delta^3(\vec{p'} - p) \\
\text{For massless fermions:} & \langle p'|p\rangle = (2\pi)^3\, 2p^0\delta^3(\vec{p'} - p)
\end{array} \tag{3.4.7}$$

Then the state number in phase space d^3p is

$$\tilde{\mathrm{d}}p = \begin{cases} \dfrac{\mathrm{d}^3p}{(2\pi)^3 2p^0} & \text{for bosons and massless fermions} \\[2ex] \dfrac{\mathrm{d}^3p}{(2\pi)^3}\dfrac{m}{p^0} & \text{for massive fermions} \end{cases}$$

Therefore, the scattering probability from unit space–time to phase space $\mathrm{d}^3p_1'\mathrm{d}^3p_2' \cdots \mathrm{d}^3p_n'$ is

$$\mathrm{d}W = \frac{[(2\pi)^4\delta^4(\sum_i p_i' - p_1 - p_2)]^2|\langle 1', 2', \ldots, n'|\hat{T}|1, 2\rangle|^2}{VT}\prod_j \tilde{\mathrm{d}}p_j'\frac{1}{S} \tag{3.4.8}$$

where S is the statistic factor $S = \prod_i n_i!$, and n_i is the number of identical particles in the final state. Since permutations of identical particles in the phase space $\mathrm{d}^3p_1'\mathrm{d}^3p_2' \cdots \mathrm{d}^3p_n'$ do not generate different quantum states, the scattering probability should be divided by the permutation number $n!$.

Using the relation

$$(2\pi)^4\,\delta^4(0) = \int \mathrm{d}^4 x \mathrm{e}^{-\mathrm{i}(p-p)\cdot x} = VT \tag{3.4.9}$$

to remove one $(2\pi)^4\delta^4(\sum_i p_i' - p_1 - p_2)$ in Eq. (3.4.8), then we get

$$\mathrm{d}W = (2\pi)^4\,\delta^4\left(\sum_i p_i' - p_1 - p_2\right)|\langle 1', 2', \ldots, n'|\,T\,|1, 2\rangle|^2 \prod_j \tilde{\mathrm{d}}p_j' \frac{1}{S} \tag{3.4.10}$$

where $\prod_j \tilde{\mathrm{d}}p_j'$ is called factor of phase space.

Let us go back to Eq. (3.4.5). If using the normalization in Eq. (3.4.7), and considering $(2\pi)^3\delta(0) = \int \mathrm{d}^3 x \mathrm{e}^{-\mathrm{i}(p-p)\cdot x} = V$, then the particle density in space of dimension 3 is

$$\begin{aligned} &\text{For bosons and massless fermions:} && \rho = 2p^0 \\ &\text{For massive fermions:} && \rho = \frac{p^0}{m} \end{aligned} \tag{3.4.11}$$

Then for the case that particles in the initial state are all bosons or massless fermions

$$\begin{aligned} \rho_1\rho_2 v_{12} &= 2p_1^0 2p_2^0\,|\vec{v}_1 - \vec{v}_2| = 4\left|\frac{\vec{p}_1}{p_1^0} - \frac{\vec{p}_2}{p_2^0}\right| p_1^0 p_2^0 \\ &= 4\left|\vec{p}_1 p_2^0 - p_1^0\vec{p}_2\right| = 4\left[(p_1\cdot p_2)^2 - m_1^2 m_2^2\right]^{\frac{1}{2}} \end{aligned} \tag{3.4.12}$$

where in the last step the condition that $\vec{p}_1$ and $\vec{p}_2$ are parallel or antiparallel is used. Finally, we can obtain the differential cross section for the scattering

$$\begin{aligned} \mathrm{d}\sigma = &\frac{1}{4[(p_1\cdot p_2)^2 - m_1^2 m_2^2]^{1/2}}|\langle 1', 2', \ldots, n'|\hat{T}|1, 2\rangle|^2 \\ &\cdot \prod_j \tilde{\mathrm{d}}p_j'\,(2\pi)^4\,\delta^4\left(\sum_{j=1}^n p_j' - p_1 - p_2\right)\frac{1}{S} \end{aligned} \tag{3.4.13}$$

Although this expression is obtained by using the condition that $\vec{p}_1$ and $\vec{p}_2$ are parallel or antiparallel, it is in the form of Lorentz invariance. Therefore, it can be used in any coordinate system.

When there are fermions in the initial state, the differential cross section can be simply obtained by making the following substitution:

$$\begin{aligned}&\int \frac{\mathrm{d}^3p}{(2\pi)^3\,2p^0} \to \int \frac{\mathrm{d}^3p}{(2\pi)^3}\frac{m}{p^0}\\ &2p^0 \to \frac{p^0}{m}\end{aligned} \tag{3.4.14}$$

in the relevant state density and phase space factor in Eq. (3.4.13).

3.4.2. *Particle lifetime and decay width*

Unstable particles decay with a certain probability. Let the probability of the particle decay per unit time be f, then at any time t, the number of particles $n(t)$ satisfies

$$\frac{\mathrm{d}n(t)}{\mathrm{d}t} = -fn(t) \tag{3.4.15}$$

Solving this equation, we get

$$n(t) = n(0)\mathrm{e}^{-ft} \tag{3.4.16}$$

At time t, the number of particles that decay in time interval $\mathrm{d}t$ is $|n'(t)|\,\mathrm{d}t$, where $n'(t) = \mathrm{d}n(t)/\mathrm{d}t$. So the mean lifetime of the particle is

$$\tau = \frac{1}{n(0)}\int_0^\infty t\,|n'(t)|\,\mathrm{d}t = \frac{1}{f} \tag{3.4.17}$$

If the wave function of the particle is $\psi(t)$, then $n(t) \propto |\psi(t)|^2$. So the wave function satisfying Eq. (3.4.16) should be

$$\psi(t) = \psi(0)\mathrm{e}^{-\mathrm{i}E_Rt-\frac{f}{2}t} \tag{3.4.18}$$

where E_R is the particle energy. Transform $\psi(t)$ into energy space by Fourier transformation

$$\begin{aligned}\chi(E) &= \int_0^\infty \psi(t)\mathrm{e}^{\mathrm{i}Et}\mathrm{d}t\\ &= \frac{C}{E_R - E - \mathrm{i}\frac{f}{2}}\end{aligned} \tag{3.4.19}$$

The distribution of particles in the energy space can be obtained by taking module square of Eq. (3.4.19). The result is a curve with a peak around

E_{R}. The width of this peak at half height is

$$\Gamma = f \tag{3.4.20}$$

which means that the decay width of the particle is equal to the probability of the particle decaying per unit time. According to the relation between the average life of particles and the decay probability in Eq. (3.4.17), it can be known that the decay width of the particle is inversely related to the mean lifetime, i.e.

$$\Gamma = 1/\tau \tag{3.4.21}$$

The probability of a particle decaying per unit space–time volume divided by the particle density is equal to the probability of a single particle decaying per unit time f. Therefore, from Eq. (3.4.20), we can obtain

$$\Gamma = \frac{\text{Decay probability per unit space–time volume}}{\text{Initial particle density}} \tag{3.4.22}$$

If the initial particle is a boson or a massless fermion, the density of the initial particle is $\rho = 2p_i^0$, therefore

$$\Gamma = \frac{1}{2p_i^0} \int \prod_{j=1}^{n} \tilde{\mathrm{d}}p_j \left| \langle 1', 2', \ldots, n' | \hat{T} | i \rangle \right|^2 (2\pi)^4 \delta^4 \left(\sum_{j=1}^{n} p_j' - p_i \right) \frac{1}{S} \tag{3.4.23}$$

For the case of massive fermion, substitution of Eq. (3.4.14) should be applied.

3.5. Kinematics of Particle Decay

3.5.1. *Two-body decay*

Consider a process of a particle with four-momentum P, mass M decaying into two particles with momentum p_1 and p_2, mass m_1 and m_2, respectively. At the center-of-mass frame $P = (M, \vec{0})$, then according to four-momentum conservation

$$P = p_1 + p_2 \tag{3.5.1}$$

We can obtain the energies of the final particles

$$E_1 = \frac{M^2 - m_2^2 + m_1^2}{2M}, \quad E_2 = \frac{M^2 - m_1^2 + m_2^2}{2M} \tag{3.5.2}$$

the momentum is

$$|p_1| = |p_2| = \frac{\sqrt{(M^2-(m_1+m_2)^2)\,(M^2-(m_1-m_2)^2)}}{2M} \tag{3.5.3}$$

For the case that the final particles are two bosons, according to the decay width formula given in Sec. 3.4, the differential decay width can be obtained

$$\mathrm{d}\Gamma = \frac{1}{32\pi^2}\,|A|^2\,\frac{|p_1|}{M^2}\mathrm{d}\Omega \tag{3.5.4}$$

where $A = \langle p_1, p_2|\hat{T}|p\rangle$ is the decay amplitude, and Ω the solid angle.

3.5.2. *Three-body decay, Dalitz diagram*

Consider a particle with four-momentum and mass (p, M) decaying into three particles with four-momenta and masses (p_1, m_1), (p_2, m_2) and (p_3, m_3). In the center-of-mass frame

$$\begin{aligned} p &= (M, \vec{0}) \\ p &= p_1 + p_2 + p_3 \end{aligned} \tag{3.5.5}$$

Denote the decay amplitude as $A = \langle p_1, p_2, p_3|\,\hat{T}\,|p\rangle$. The decay rate is proportional to $|A|^2$ and phase space factor, i.e.

$$W \propto \int |A|^2 \frac{\mathrm{d}^3p_1}{E_1}\frac{\mathrm{d}^3p_2}{E_2}\frac{\mathrm{d}^3p_3}{E_3}\delta^4\left(p_1+p_2+p_3-p\right) \tag{3.5.6}$$

Using the relation

$$\frac{1}{2E} = \int \mathrm{d}E\delta(p^2-m^2)\theta(E) \tag{3.5.7}$$

we can change the integral of d^3p_3 to integral of four-momentum space

$$\begin{aligned} W &\propto 2\int |A|^2 \frac{\mathrm{d}^3p_1}{E_1}\frac{\mathrm{d}^3p_2}{E_2}\mathrm{d}^4p_3\delta\left(p_3^2-m_3^2\right)\delta^4\left(p_1+p_2+p_3-p\right)\theta(E_3) \\ &= 2\int \frac{\mathrm{d}^3p_1}{E_1}\frac{\mathrm{d}^3p_2}{E_2}\delta\left((p-p_1-p_2)^2-m_3^2\right)\theta(E_3)\,|A|^2 \end{aligned} \tag{3.5.8}$$

Integrating for the azimuthal angles of momenta $\vec{p}_1$ and $\vec{p}_2$ in Eq. (3.5.8), we obtain

$$W \propto 8\pi^2 \int \mathrm{d}E_1\mathrm{d}E_2\,|A|^2 \tag{3.5.9}$$

Because the kinetic energy of the particle is the total energy minus the static energy, i.e.

$$T_1 = E_1 - m_1, \quad T_2 = E_2 - m_2, \quad T_3 = E_3 - m_3 \tag{3.5.10}$$

the decay probability can also be expressed as the integral of kinetic energies

$$W \propto 8\pi^2 \int \mathrm{d}T_1 \mathrm{d}T_2 \, |A|^2 \tag{3.5.11}$$

In addition, define that the sum of kinetic energies of particles in the final state minus the sum of kinetic energy of particle in the initial state as Q, then

$$\begin{aligned} Q &= T_1 + T_2 + T_3 - T \\ &= E_1 + E_2 + E_3 - m_1 - m_2 - m_3 \\ &= M - m_1 - m_2 - m_3 \end{aligned} \tag{3.5.12}$$

where T is the kinetic energy of particle in the initial state, and $T = 0$ in the rest frame of the initial particle. Equation (3.5.12) shows that the sum of the kinetic energies of the particles in final state is a constant at the center-of-mass frame. From plane geometry, we know that the sum of the lengths of the perpendicular lines from any point in an equilateral triangle to three sides is also a constant, so we can use this property to represent the kinetic energy distribution of the final particles as points in an equilateral triangle. We can draw an equilateral triangle with its height being Q, as shown in Fig. 3.3. Then any possible momentum distribution can be represented by a point in this triangle. Such diagram is called Dalitz diagram [34]. Dalitz diagram can also be defined by other kinematic variables, as long as these kinematic variables add up to be a constant.

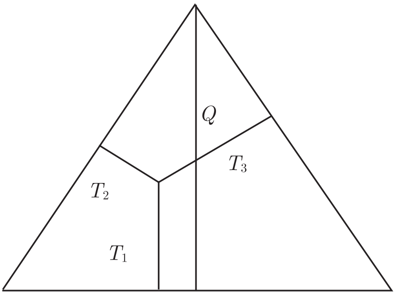


Fig. 3.3. Dalitz diagram.

Not every point in the triangle represents a physically possible kinematic distribution. The vertex of the triangle obviously does not work, because it does not satisfy the conservation of four-momentum. According to momentum conservation $\vec{p}_3 = -(\vec{p}_1 + \vec{p}_2)$, then

$$p_3^2 = p_1^2 + p_2^2 + 2p_1 p_2 \cos\alpha, \quad |\cos\alpha| \le 1 \tag{3.5.13}$$

we can obtain

$$4p_1^2 p_2^2 \ge (p_3^2 - p_1^2 - p_2^2)^2 \tag{3.5.14}$$

in the above two equations, p_i is the module of the three-momentum. Let us consider the condition that all the masses of the three particles are the same, $m_1 = m_2 = m_3 = m$. Then in the non-relativistic case, the kinetic energy of the particle is $T_i = p_i^2/2m$. Dividing both sides of Eq. (3.5.14) by $(2m)^2$, we get

$$4T_1 T_2 \ge (T_3 - T_1 - T_2)^2 \tag{3.5.15}$$

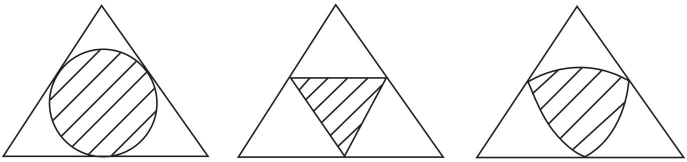
Equation (3.5.15) corresponds to boundary condition

$$4T_1 T_2 = (T_3 - T_1 - T_2)^2 \tag{3.5.16}$$

This represents the inner circle of the triangle, as shown in Fig. 3.4(a) [35].

In the condition of extreme relativistic $T_i = E_i = p_i$, one can obtain from Eq. (3.5.14) that the boundary condition becomes

$$T_3 = \begin{cases} T_1 + T_2 \quad \xleftarrow{p_3} \begin{array}{l} \longrightarrow p_1 \\ \longrightarrow p_2 \end{array} \\ T_1 - T_2 \quad \xleftarrow{p_1} \begin{array}{l} \longrightarrow p_2 \\ \longrightarrow p_3 \end{array} \\ -T_1 + T_2 \xleftarrow{p_2} \begin{array}{l} \longrightarrow p_3 \\ \longrightarrow p_1 \end{array} \end{cases} \tag{3.5.17}$$



(a) (b) (c)

Fig. 3.4. Physical region of three-body decays in Dalitz diagram: (a) non-relativistic condition; (b) extreme relativistic condition; (c) ordinary condition.

This represents a triangle, as shown in Fig. 3.4(b). For ordinary condition, the physical region would be somewhere between the non-relativistic and the extreme relativistic case, which is shown in Fig. 3.4(c) [35].

It is not difficult to prove that the area element in a triangle $\mathrm{d}S \propto \mathrm{d}T_1\mathrm{d}T_2$. Then the decay rate

$$W \propto |A|^2 \mathrm{d}T_1 \mathrm{d}T_2 \propto |A|^2 \mathrm{d}S \tag{3.5.18}$$

It can be known from Eq. (3.5.18) that if experimental data are marked in the triangle, then the density of the event points is proportional to the decay amplitude squared $|A|^2$. Therefore, the distribution of event points in the triangle can reflect the dynamics of the decay. It is possible to determine the spin and parity of the initial particle by analyzing the distribution of experimental points in Dalitz diagram.

The following is a brief introduction about how to determine the spin of K^+ meson using Dalitz diagram.

The process $\mathrm{K}^+ \to \pi^+\pi^+\pi^-$ is used. The momenta of these three π mesons are marked in Dalitz diagram as shown in Fig. 3.5.

Analysis: Let the orbital angular momentum between two π^+ mesons be L. These two π^+'s are identical bosons, so $L = 0, 2, 4, \ldots$ must be even numbers. In addition, let the orbital angular momentum of the π^- meson relevant to $\pi^+\pi^+$ pair be l. The value of l can be $0, 1, 2, \ldots$.

(1) If $L \geq 2$, that is, the relative orbital angular momentum of $\pi^+\pi^+$ pair is larger, so the kinetic energy of $\pi^+\pi^+$ should be larger, then the momentum of π^- would be smaller. If this is the case, the distribution of event points in the upper part of the circle in the Dalitz diagram will decrease.

(2) If $l \geq 1$, the orbital angular momentum of π^- relevant to $\pi^+\pi^+$ is larger, then its kinetic energy is also larger. So the distribution of event points in the lower part of the circle in the Dalitz diagram will decrease.

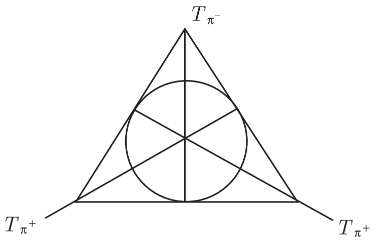


Fig. 3.5.

Analyzing the distribution of the event points measured in experiment, no deviation was found in the distribution. The distribution of event points is uniform. Therefore, the above two possibilities are excluded. L and l can only be $L = 0$ and $l = 0$. Because the spin of pion has been known to be 0, the spin of K^+ can only be $J = |L + l|, |L + l - 1|, \ldots, |L - l|$. So it can be determined that the spin of K^+ is 0.

References

1. Anderson C D, Neddermeyer S H. *Phys. Rev.*, 1936, 50: 63.
2. Neddermeyer S H, Anderson C D. *Phys. Rev.*, 1937, 51: 884.
3. Neddermeyer S H, Anderson C D. *Phys. Rev.*, 1938, 54: 88.
4. Neddermeyer S H, Anderson C D. *Rev. Mod. Phys.*, 1939, 11: 191.
5. Conversi M, Pancini E, Piccioni O. *Phys. Rev.*, 1947, 71: 209.
6. Reines F, Cowan C L. *Phys. Rev.*, 1953, 90: 492.
7. Reines F, Cowan C L. *Phys. Rev.*, 1953, 92: 830.
8. Cowan C L, *et al. Science*, 1956, 124: 103.
9. Danby G, *et al. Phys. Rev. Lett.*, 1962, 9: 36.
10. Kodama K, *et al.* (DONUT Collab.) *Phys. Lett.*, 2001, B504: 218.
11. Goldhaber M, Grodzins L, Sunyar A W. *Phys. Rev.*, 1958, 109: 1015.
12. Yukawa H. *Proc. Phys. Math. Soc. Jap.*, 1935, 17: 48–57.
13. Latte C M G, Muirhead H, Occhialini G P S, *et al. Nature*, 1947, 159: 694.
14. Durbin R, Loar H, Steinberger. *J. Phys. Rev.*, 1951, 83: 646.
15. Durbin R, Loar H, Steinberger. *J. Phys. Rev.*, 1951, 84: 581.
16. Clark D. L, Roberts A, Wilson R. *Phys. Rev.*, 1952, 85: 523.
17. Cartwright W F, *et al. Phys. Rev.*, 1953, 91: 677.
18. Plano R, *et al. Phys. Rev. Lett.*, 1959, 3: 525.
19. Chamberlain O, *et al. Phys. Rev.*, 1955, 100: 947.
20. Cork B, *et al. Phys. Rev.*, 1957, 104: 1193.
21. Rochester G D, Butler C C. *Nature*, 1947, 160: 855.
22. Pais A. *Phys. Rev.*, 1952, 86: 663–672.
23. Gell-Mann M. *Phys. Rev.*, 1953, 92: 833–834.
24. Nakano T, Nishijima K. *Prog. Theor. Phys.*, 1953, 10: 581.
25. Anderson H L, *et al. Phys. Rev.*, 1952, 85: 934.
26. Anderson H L, *et al. Phys. Rev.*, 1952, 85: 936.
27. Lindenbaum S J, Yuan L C L. *Phys. Rev.*, 1953, 92: 1578.
28. Ashkin J, *et al. Phys. Rev.*, 1954, 93: 1129.
29. Stonehill D, *et al. Phys. Rev. Lett.*, 1961, 6: 624.
30. Erwin A R, *et al. Phys. Rev. Lett.*, 1961, 6: 628.
31. Maglic B C, *et al. Phys. Rev. Lett.*, 1961, 7: 178.
32. Pevsner A, *et al. Phys. Rev. Lett.*, 1961, 7: 421.
33. Alff-Steinberger C, Berley D, Colley D. *Phys. Rev. Lett.*, 1962, 9: 322.
34. Dalitz R H. *Phys. Rev.*, 1954, 94: 1046.
35. Lee T D. *Introduction to Particle Physics and Field Theory.* Science Press, 1980.

This page intentionally left blank

Chapter 4

Isospin and G Parity

4.1. Isospin

4.1.1. *The concept of isospin*

Historically, the concept of isospin was proposed by considering the similarity between protons and neutrons under strong interaction. The study of the nucleus shows that strong interactions between proton–proton, proton–neuron and neutron–neutron are almost of no difference. There are very similar energy-level structures and spins for mirror nuclei. A mirror nucleus is a new nucleus made by swapping the number of protons and neutrons in nucleus. For example, there are three protons and four neutrons in Li^7 nucleus, while four protons and three neutrons in Be^7. Such two nuclei are mirror nuclei to each other. The slight difference between the structures of energy levels of mirror nuclei can be attributed to difference in electromagnetic interactions between protons and neutrons. There are sufficient reasons to believe that strong interaction is independent of the electric charge of the particle. This is called charge independence of strong interaction.

If protons and neutrons in a nuclear system are exchanged, the property of the whole system almost does not change. In 1932, according to such symmetry between protons and neutrons in strong interaction in nuclear system, Heisenberg proposed the concept of isospin [1]. He drew an analogy of proton and neutron to two spin states of electron $\mathrm{S} = \pm 1/2$: there are two spin states for electron, which are relevant to the third component of spin $+1/2$ and $-1/2$. They are two different states for the same particle. The analogy is that proton and neutron are also two different states of one particle. Spin is vector, which has three components. Assume that the quantity describing different states for proton and neutron is also a vector, which is called isospin. The third component of isospin can take values of $+1/2$ and

$-1/2$, whose eigenstates are relevant to proton and neutron, respectively. In the presence of the electromagnetic field, the energy levels of an electron for two different spin states split, which leads to difference between these two states. Similarly, in the presence of the electromagnetic field, the behaviors of proton and neutron also tend to be different, and the energy levels of mirror nuclei are slightly different. When the external electromagnetic field is removed, the splitting energy levels of electrons recombine and revert to degeneracy. If there is no electromagnetic interaction in a nuclear system, the mirror nucleus will have the same energy-level structure, and there will be a symmetry under the exchange of proton and neutron in the strong interactions. Certainly, the different behavior of proton and neutron in strong interaction may also come from the slight difference between the masses of proton and neutron, which will not be discussed in detail here.

The spin of electron $\vec{S}$ is a vector in coordinate space, while isospin is only a vector of an abstract space, which is denoted as $\vec{I}$. Such vector is used to tag the states of nucleon. Therefore, the space of isospin is a space of particle states. There are only two states for nucleon, proton and neutron. We can use a mathematical tool describing the spin states of electrons to describe the isospin of nucleon. Assume that proton corresponds to the state with the third component of isospin $I_3 = 1/2$, and neutron corresponds to the isospin state $I_3 = -1/2$. Totally, nucleon belongs to a doublet of isospin being 1/2, which can be represented by a 2×1 column matrix

$$\begin{pmatrix} \psi_{\mathrm{p}} \\ \psi_{\mathrm{n}} \end{pmatrix} \tag{4.1.1}$$

Then the states for proton and neutron are

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{4.1.2}$$

respectively.

In this representation, the components of the operator of isospin $\vec{\tau}/2$ are proportional to Pauli matrices

$$\tau_1 = \sigma_1, \quad \tau_2 = \sigma_2, \quad \tau_3 = \sigma_3 \tag{4.1.3}$$

where σ_i's are Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{4.1.4}$$

Sometimes, we also use

$$\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3) \tag{4.1.5}$$

to represent the three Pauli matrices as a whole.

$\begin{pmatrix}1\\0\end{pmatrix}$ and $\begin{pmatrix}0\\1\end{pmatrix}$ are all eigen states of $\tau_3/2$, and the eigenvalues are $1/2$ and $-1/2$,

$$\begin{aligned}\frac{1}{2}\sigma_3\begin{pmatrix}1\\0\end{pmatrix} &= \frac{1}{2}\begin{pmatrix}1 & 0\\0 & -1\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1\\0\end{pmatrix}\\ \frac{1}{2}\sigma_3\begin{pmatrix}0\\1\end{pmatrix} &= \frac{1}{2}\begin{pmatrix}1 & 0\\0 & -1\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} = -\frac{1}{2}\begin{pmatrix}0\\1\end{pmatrix}\end{aligned} \tag{4.1.6}$$

and the isospin squared operator is

$$\left(\frac{\vec{\tau}}{2}\right)^2 = \left(\frac{1}{2}\sigma_1\right)^2 + \left(\frac{1}{2}\sigma_2\right)^2 + \left(\frac{1}{2}\sigma_3\right)^2 = \frac{3}{4} = \frac{1}{2}\left(\frac{1}{2} + 1\right) \tag{4.1.7}$$

therefore, $\vec{\tau}^2$ commutes with all τ_1, τ_2 and τ_3.

$\begin{pmatrix}1\\0\end{pmatrix}$ and $\begin{pmatrix}0\\1\end{pmatrix}$ are also eigenstates of $(\vec{\tau}/2)^2$, and the relevant eigenvalue is $I(I+1) = 3/4$, i.e., $I = 1/2$. So, $\begin{pmatrix}1\\0\end{pmatrix}$ and $\begin{pmatrix}0\\1\end{pmatrix}$ are the common eigenstate of $(\vec{\tau}/2)^2$ and $\tau_3/2$ with eigenvalues $I = 1/2$, $I_3 = 1/2$ and $-1/2$, respectively. Usually, $|I, I_3\rangle$ is used to denote the eigenstate of total isospin I, and the third component I_3.

The rule for the superposition of two isospin vectors is the same as that for the addition of angular momenta. The coupling of the isospin of two nucleons is

$$\begin{aligned}&\text{pp:} \quad \left|\frac{1}{2},\frac{1}{2}\right\rangle\left|\frac{1}{2},\frac{1}{2}\right\rangle = \left|1,1\right\rangle\\ &\text{pn:} \quad \left|\frac{1}{2},\frac{1}{2}\right\rangle\left|\frac{1}{2},-\frac{1}{2}\right\rangle = \frac{1}{\sqrt{2}}\left|1,0\right\rangle + \frac{1}{\sqrt{2}}\left|0,0\right\rangle\\ &\text{np:} \quad \left|\frac{1}{2},-\frac{1}{2}\right\rangle\left|\frac{1}{2},\frac{1}{2}\right\rangle = \frac{1}{\sqrt{2}}\left|1,0\right\rangle - \frac{1}{\sqrt{2}}\left|0,0\right\rangle\\ &\text{nn:} \quad \left|\frac{1}{2},-\frac{1}{2}\right\rangle\left|\frac{1}{2},-\frac{1}{2}\right\rangle = \left|1,-1\right\rangle\end{aligned} \tag{4.1.8}$$

where the coefficients of the coupled isospin vectors in the right-hand sides of the above equations are Clebsch–Gordan coefficients, which can be obtained by referring to the table of Clebsch–Gordan coefficients.

From Eq. (4.1.8), the total isospin state $|1,1\rangle$, $|1,0\rangle$, $|1,-1\rangle$ and $|0,0\rangle$ can also be expressed in terms of the uncoupled representations of isospin wave function for nucleons,

$$\begin{aligned}
|1,1\rangle &= \left|\frac{1}{2},\frac{1}{2}\right\rangle \left|\frac{1}{2},\frac{1}{2}\right\rangle \\
|1,0\rangle &= \frac{1}{\sqrt{2}}\left[\left|\frac{1}{2},\frac{1}{2}\right\rangle \left|\frac{1}{2},-\frac{1}{2}\right\rangle + \left|\frac{1}{2},-\frac{1}{2}\right\rangle \left|\frac{1}{2},\frac{1}{2}\right\rangle\right] \\
|1,-1\rangle &= \left|\frac{1}{2},-\frac{1}{2}\right\rangle \left|\frac{1}{2},-\frac{1}{2}\right\rangle \\
|0,0\rangle &= \frac{1}{\sqrt{2}}\left[\left|\frac{1}{2},\frac{1}{2}\right\rangle \left|\frac{1}{2},-\frac{1}{2}\right\rangle - \left|\frac{1}{2},-\frac{1}{2}\right\rangle \left|\frac{1}{2},\frac{1}{2}\right\rangle\right]
\end{aligned} \tag{4.1.9}$$

Equation (4.1.9) indicates that the three states with total isospin 1 are symmetric states relative to the transformation of two-nucleon permutation, and the state with total isospin 0 is a singlet state, which is antisymmetric under nucleon permutation transformation. This is the same as that when two angular momenta of electrons are coupled.

4.1.2. *Isospin transformation*

There is a symmetry in strong interaction under the exchange of protons and neutrons, that is, when proton is replaced by neutron or neutron replaced by proton, the status of strong interaction does not change. But this is just a special case of the charge-independent symmetry of the strong interaction. More generally, on matter, protons are replaced by neutrons, or by a mixture of protons and neutrons superimposed in a certain proportion, the strong interactions of the system do not change. So strong interaction is invariant under any transformation between p and n. Define the transformation of p and n as

$$\begin{pmatrix} \mathrm{p}' \\ \mathrm{n}' \end{pmatrix} = U \begin{pmatrix} \mathrm{p} \\ \mathrm{n} \end{pmatrix} \tag{4.1.10}$$

where U is a complex matrix of 2×2. To keep probability conservation, U should be unitary matrix, which satisfies $U^{+}U = 1$. The collection of these U transformation matrices makes up U(2) group, and these 2×2 unitary matrices constitute self-representation of the U(2) group. The modules of these matrices are 1. Therefore, by removing one factor irrelevant to the transformation between p and n, the determinants of these matrices can be made to be 1. A new group can be obtained in this way, which is the SU(2) group. Then we know that the simplest complex matrix group describing the transformation between p and n with probability conserved is the SU(2) group. The transformation under the SU(2) group is called isospin transformation.

To show the transformation of the nucleon field $\psi(x)$ under isospin transformation, we substitute the nucleon states p and n in Eq. (4.1.10) with their relevant field operators $\psi_{\mathrm{p}}(x)$ and $\psi_{\mathrm{n}}(x)$, and U matrix is changed to element of SU(2) group, denoted as S. Then Eq. (4.4.10) becomes

$$\begin{pmatrix} \psi'_{\mathrm{p}}(x) \\ \psi'_{\mathrm{n}}(x) \end{pmatrix} = S \begin{pmatrix} \psi_{\mathrm{p}}(x) \\ \psi_{\mathrm{n}}(x) \end{pmatrix} \tag{4.1.11}$$

For each matrix S, there is a relevant unitary operator in Hilbert space. Then Eq. (4.1.11) can be written as

$$\hat{S} \begin{pmatrix} \psi_{\mathrm{p}}(x) \\ \psi_{\mathrm{n}}(x) \end{pmatrix} \hat{S}^{+} = S \begin{pmatrix} \psi_{\mathrm{p}}(x) \\ \psi_{\mathrm{n}}(x) \end{pmatrix} \tag{4.1.12}$$

The creation and annihilation operators for p, n and p̄, n̄ can be defined through the Fourier transformation of $\psi_{\mathrm{p}}(x)$ and $\psi_{\mathrm{n}}(x)$

$$\begin{aligned} \psi_{\mathrm{p}}(x) &= \int \frac{\mathrm{d}^3 k}{(2\pi)^3} \frac{m}{k^0} \sum_{\alpha=1,2} \left[a_{\alpha}^{\mathrm{p}}(k) u^{(\alpha)}(k) e^{-\mathrm{i}k\cdot x} + b_{\alpha}^{\mathrm{p}^+}(k) v^{(\alpha)}(k) e^{\mathrm{i}k\cdot x} \right] \\ \psi_{\mathrm{n}}(x) &= \int \frac{\mathrm{d}^3 k}{(2\pi)^3} \frac{m}{k^0} \sum_{\alpha=1,2} \left[a_{\alpha}^{\mathrm{n}}(k) u^{(\alpha)}(k) e^{-\mathrm{i}k\cdot x} + b_{\alpha}^{\mathrm{n}^+}(k) v^{(\alpha)}(k) e^{\mathrm{i}k\cdot x} \right] \end{aligned} \tag{4.1.13}$$

where $a_{\alpha}^{\mathrm{p}}(k)$, $b_{\alpha}^{\mathrm{p}^+}(k)$ and $a_{\alpha}^{\mathrm{n}}(k)$, $b_{\alpha}^{\mathrm{n}^+}(k)$ are the annihilation and creation operators of p, p̄ and n, n̄, respectively.

Substitute Eq. (4.1.13) into (4.1.11), and compare the left and right sides, then we can obtain

$$\hat{S} \begin{pmatrix} a_{\alpha}^{\mathrm{p}}(k) \\ a_{\alpha}^{\mathrm{n}}(k) \end{pmatrix} \hat{S}^{+} = S \begin{pmatrix} a_{\alpha}^{\mathrm{p}}(k) \\ a_{\alpha}^{\mathrm{n}}(k) \end{pmatrix} \tag{4.1.14}$$

and

$$\hat{S}\begin{pmatrix} b_\alpha^{\mathrm{p}^+}(k) \\ b_\alpha^{\mathrm{n}^+}(k) \end{pmatrix}\hat{S}^+ = S\begin{pmatrix} b_\alpha^{\mathrm{p}^+}(k) \\ b_\alpha^{\mathrm{n}^+}(k) \end{pmatrix} \tag{4.1.15}$$

The above are the isospin transformations of the annihilation operator of nucleon and creation operator of antinucleon. Next, let's derive the transformation of the creation operator of nucleon and annihilation operator of antinucleon. Find the Hermitian conjugate of both sides of Eq. (4.1.14), then we can get

$$\left(\hat{S}a_\alpha^{\mathrm{p}^+}(k)\hat{S}^+, \hat{S}a_\alpha^{\mathrm{n}^+}(k)\hat{S}^+\right) = \left(a_\alpha^{\mathrm{p}^+}(k), a_\alpha^{\mathrm{n}^+}(k)\right)S^+ \tag{4.1.16}$$

Write S explicitly as

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \quad 和 \quad S^+ = \begin{pmatrix} S_{11}^* & S_{21}^* \\ S_{12}^* & S_{22}^* \end{pmatrix} \tag{4.1.17}$$

Thus, Eq. (4.1.16) can be further written as

$$\begin{aligned} &\left(\hat{S}a_\alpha^{\mathrm{p}^+}(k)\hat{S^+}, \hat{S}a_\alpha^{\mathrm{n}^+}(k)\hat{S}^+\right) \\ &= \left(a_\alpha^{\mathrm{p}^+}(k)S_{11}^* + a_\alpha^{\mathrm{n}^+}(k)S_{12}^*, a_\alpha^{\mathrm{p}^+}(k)S_{21}^* + a_\alpha^{\mathrm{n}^+}(k)S_{22}^*\right) \end{aligned} \tag{4.1.18}$$

Keeping each element in the above equation equal, Eq. (4.1.18) can be written as a column matrix again

$$\hat{S}\begin{pmatrix} a_\alpha^{\mathrm{p}^+}(k) \\ a_\alpha^{\mathrm{n}^+}(k) \end{pmatrix}\hat{S}^+ = S^*\begin{pmatrix} a_\alpha^{\mathrm{p}^+}(k) \\ a_\alpha^{\mathrm{n}^+}(k) \end{pmatrix} \tag{4.1.19}$$

From Eq. (4.1.15), one can similarly obtain

$$\hat{S}\begin{pmatrix} b_\alpha^{\mathrm{p}}(k) \\ b_\alpha^{\mathrm{n}}(k) \end{pmatrix}\hat{S}^+ = S^*\begin{pmatrix} b_\alpha^{\mathrm{p}}(k) \\ b_\alpha^{\mathrm{n}}(k) \end{pmatrix} \tag{4.1.20}$$

Comparing the above two equations and Eqs. (4.1.14) and (4.1.15), it implies that the isospin transformations of the creation and annihilation operators of nucleon and antinucleon are complex conjugate of each other.

Define the nucleon states for p and n as

$$\begin{aligned}|\mathrm{p}\rangle &= a_\alpha^{\mathrm{p}^+}(k)\,|0\rangle, \quad |\bar{\mathrm{p}}\rangle = b_\alpha^{\mathrm{p}^+}(k)\,|0\rangle \\ |\mathrm{n}\rangle &= a_\alpha^{\mathrm{n}^+}(k)\,|0\rangle, \quad |\bar{\mathrm{n}}\rangle = b_\alpha^{\mathrm{n}^+}(k)\,|0\rangle\end{aligned} \tag{4.1.21}$$

and assume that the vacuum state is invariant under isospin transformation, i.e.,

$$\hat{S}\,|0\rangle = |0\rangle \tag{4.1.22}$$

Then we can obtain the isospin transformation of nucleon and antinucleon from Eqs. (4.1.15) and (4.1.19)

$$\hat{S}\begin{pmatrix}|\bar{\mathrm{p}}\rangle \\ |\bar{\mathrm{n}}\rangle\end{pmatrix} = S\begin{pmatrix}|\bar{\mathrm{p}}\rangle \\ |\bar{\mathrm{n}}\rangle\end{pmatrix} \tag{4.1.23}$$

and

$$\hat{S}\begin{pmatrix}|\mathrm{p}\rangle \\ |\mathrm{n}\rangle\end{pmatrix} = S^*\begin{pmatrix}|\mathrm{p}\rangle \\ |\mathrm{n}\rangle\end{pmatrix} \tag{4.1.24}$$

Note that Eqs. (4.1.23) and (4.1.24) indicate that the isospin transformations of the nucleon and the antinucleon are complex conjugates of each other.

The isospin transformation matrix S is the element of SU(2) group. Its self-representation is

$$S = e^{-i\frac{1}{2}\vec{\sigma}\cdot\vec{\theta}} \tag{4.1.25}$$

where $\vec{\theta} = (\theta_1,\ \theta_2,\ \theta_3)$ are SU(2) group parameters. Expand Eq. (4.1.25) as

$$S = \sum_n \frac{1}{n!}\left(-i\frac{1}{2}\vec{\sigma}\cdot\vec{\theta}\right)^n \tag{4.1.26}$$

With $\sigma_i^2 = 1$ and the anti-commutation property of Pauli matrices, we can get, for any positive integer m,

$$\left(-i\frac{1}{2}\vec{\sigma}\cdot\vec{\theta}\right)^{2m} = (-1)^m\left(\frac{\left|\vec{\theta}\right|}{2}\right)^{2m} \tag{4.1.27}$$

and

$$\left(-\mathrm{i}\frac{1}{2}\vec{\sigma}\cdot\vec{\theta}\right)^{2m+1}=-\mathrm{i}\frac{\vec{\sigma}\cdot\vec{\theta}}{\left|\vec{\theta}\right|}(-1)^m\left(\frac{\left|\vec{\theta}\right|}{2}\right)^{2m+1} \tag{4.1.28}$$

where $\left|\vec{\theta}\right|=\sqrt{\theta_1^2+\theta_2^2+\theta_3^2}$. According to Eqs. (4.1.26)–(4.1.28), another expression for the isospin transformation matrix can be obtained as

$$S=\cos\frac{\left|\vec{\theta}\right|}{2}-\mathrm{i}\frac{\vec{\sigma}\cdot\vec{\theta}}{\left|\vec{\theta}\right|}\sin\frac{\left|\vec{\theta}\right|}{2} \tag{4.1.29}$$

Considering Eq. (4.1.29), the following relation can be obtained:

$$\sigma_2 S=S^*\sigma_2 \tag{4.1.30}$$

Multiply σ_2 by left on both sides of Eq. (4.1.23), and note that σ_2 as a matrix can commute with the operator $\hat{S}$ in Hilbert space, then one can get

$$\hat{S}\sigma_2\begin{pmatrix}|\bar{\mathrm{p}}\rangle\\|\bar{\mathrm{n}}\rangle\end{pmatrix}=\sigma_2 S\begin{pmatrix}|\bar{\mathrm{p}}\rangle\\|\bar{\mathrm{n}}\rangle\end{pmatrix} \tag{4.1.31}$$

In addition, using the relation in Eq. (4.1.30), Eq. (4.1.31) becomes

$$\hat{S}\begin{pmatrix}|\bar{\mathrm{n}}\rangle\\-|\bar{\mathrm{p}}\rangle\end{pmatrix}=S^*\begin{pmatrix}|\bar{\mathrm{n}}\rangle\\-|\bar{\mathrm{p}}\rangle\end{pmatrix} \tag{4.1.32}$$

Comparing Eqs. (4.1.32) and (4.1.24), we can know that the doublet $\begin{pmatrix}|\bar{\mathrm{n}}\rangle\\-|\bar{\mathrm{p}}\rangle\end{pmatrix}$ and $\begin{pmatrix}|p\rangle\\|n\rangle\end{pmatrix}$ behave the same way under isospin transformation. That says, if viewing proton p as isospin state upwards ↑, and neutron the state downwards ↓, then the antineutron will be the isospin upward state ↑, which is the same state as p under isospin transformation; while it is $-\bar{p}$, not $\bar{p}$, the isospin downward state ↓. $-\bar{p}$ and n behave the same way under isospin transformation. As we will see in Chapter 5, such property about isospin can be extended to quark system $\begin{pmatrix}u\\d\end{pmatrix}$, whose isospin is also 1/2. Just because of this property, the isospin wave function of π^0 is $\frac{\mathrm{d}\bar{\mathrm{d}}-\mathrm{u}\bar{\mathrm{u}}}{\sqrt{2}}$, instead of $\frac{\mathrm{d}\bar{\mathrm{d}}+\mathrm{u}\bar{\mathrm{u}}}{\sqrt{2}}$.

4.1.3. *The law of isospin conservation in strong interaction*

For charge independence of strong interaction, a more precise mathematical expression is strong interaction being invariant under isospin transformation of the SU(2) group. Only the components of isospin change under SU(2) transformation, and the total isospin of the system will not change. Therefore, strong interaction can only depend on the total isospin I, instead of depending on the components of isospin. Otherwise, strong interaction cannot be invariant under SU(2) group transformation.

The hamiltonian of a strong interacting system can be denoted as $\hat{H}_{\mathrm{st}}$, then the invariance of $\hat{H}_{\mathrm{st}}$ under isospin transformation is

$$\hat{S}\hat{H}_{\mathrm{st}}\hat{S}^{+} = \hat{H}_{\mathrm{st}} \tag{4.1.33}$$

Next, we shall prove that both the total isospin and its components are invariant under strong interactions.

Proof. For an infinitesimal isospin transformation, $S = 1 - \frac{\mathrm{i}}{2}\vec{\sigma} \cdot \vec{\varepsilon}$, where $\vec{\varepsilon} = (\varepsilon_1, \varepsilon_2, \varepsilon_3)$ are infinitesimal transformation parameters, and ε_i's are real numbers. There is the relevant transformation operator in Hilbert space $\hat{S} = 1 - i\hat{\vec{I}} \cdot \vec{\varepsilon}$, where $\hat{\vec{I}}$ is called the isospin operator. □

According to the unitarity of the operator $\hat{S}$

$$\hat{S}\hat{S}^{+} = \left(1 - \mathrm{i}\hat{\vec{I}} \cdot \vec{\varepsilon}\right)\left(1 + \mathrm{i}\hat{\vec{I}}^{+} \cdot \vec{\varepsilon}\right) = 1 \tag{4.1.34}$$

we can obtain

$$\left(\hat{\vec{I}} - \hat{\vec{I}}^{+}\right) \cdot \vec{\varepsilon} + O(\varepsilon) = 0 \tag{4.1.35}$$

Ignoring higher order terms, then

$$\hat{\vec{I}} = \hat{\vec{I}}^{+} \tag{4.1.36}$$

That is to say, the isospin operator $\hat{\vec{I}}$ is Hermitian.

Substituting $\hat{S} = 1 - \mathrm{i}\hat{\vec{I}} \cdot \vec{\varepsilon}$ into Eq. (4.1.33), and considering the Hermiticity of $\hat{\vec{I}}$, we can obtain

$$\left(1 - \mathrm{i}\hat{\vec{I}} \cdot \vec{\varepsilon}\right)\hat{H}_{\mathrm{st}}\left(1 + \mathrm{i}\hat{\vec{I}} \cdot \vec{\varepsilon}\right) = \hat{H}_{\mathrm{st}} \tag{4.1.37}$$

From Eq. (4.1.37), we get

$$\left[\hat{H}_{\mathrm{st}}, \hat{\vec{I}}\right] = 0 \tag{4.1.38}$$

Equation (4.1.38) shows that the isospin operator $\hat{\vec{I}}$ commute with the Hamiltonian of strong interaction. Therefore, the isospin vector is conserved in strong interaction, i.e., both the quantum numbers of the total isospin I and the third component I_3 are invariant in strong interaction. This is the theorem of isospin conservation.

Experiments shows that electromagnetic interaction violates the total value of isospin, but the third component I_3 is conserved, while both I and I_3 are not conserved in weak interaction.

4.1.4. *The isospin of mesons and baryons*

An important principle in determining the isospin of particles is to see the behavior of particles in strong interaction. First, find out how many particles exist with similar masses. Such collection of particles will be potentially to form an isospin multiplet. A multiplet with an isotopic value of I contains a number of particles of $2I+1$, similar to the case of angular momentum. For example, p and n constitute a doublet state with an isospin of 1/2, and particles with isospin 1 form a triplet, the third components of which are 1, 0, −1, respectively. Secondly, we should see whether the isospin quantum number thus determined is conserved under strong interaction. The isospin quantum numbers of some mesons and baryons are discussed in what follows:

4.1.4.1. *Nucleons p and n*

As already known, p and n compose isotropic doublet of isospin 1/2, the third components of which I_3 are 1/2 and −1/2, respectively.

Note that the charge of nucleon Q is related to the third component of its isospin I_3 and the baryon number B as

$$Q = I_3 + B/2 \tag{4.1.39}$$

4.1.4.2. *Isospin of* π *meson*

There are three charge states for pion: π^+, π^0 and π^-, the masses of which are $m_{\pi^\pm} = 139.6$ MeV and $m_{\pi^0} = 135.0$ MeV. Their masses are very similar, both are about 140 MeV. So, (π^+, π^0, π^-) can be composed into a triplet, of which the isospin is $I = 1$, and their third component of isospin is $I_3 = 1, 0, -1$, respectively.

Then let's consider the nuclear reactions that pions participate, so that we can check whether the isospin determined in this way is being conserved in strong interaction. Let's see the following reaction:

$$\mathrm{p} + \mathrm{p} \to \pi^+ + \mathrm{d} \tag{4.1.40}$$

$$\mathrm{p} + \mathrm{n} \to \pi^0 + \mathrm{d} \tag{4.1.41}$$

where deuteron d is composed of p and n, the isospin of which can be $I_d = 1$ or 0. There are no other bound states of nucleons, such as pp or nn, that can make up an isospin triplet with deuteron d, so the isospin of deuteron cannot be 1, it can only be $I_d = 0$.

The lefthand side of Eq. (4.1.40) is the isospin state of $|1, 1\rangle$. Since $I_d = 0$, and π^+ is the isospin state $|1, 1\rangle$, the righthand side of Eq. (4.1.40) is also an isospin state of $|1, 1\rangle$. This conforms to the law of isospin conservation. It can be inferred from the charge conjugated reaction of Eq. (4.1.40) that the isospin state of π^- being $|1, -1\rangle$ is also reasonable.

For Eq. (4.1.41), the isospin state formed by p and n on the left-hand side is

$$|pn\rangle = \frac{1}{\sqrt{2}} \left[|1, 0\rangle + |0, 0\rangle \right] \tag{4.1.42}$$

For the right-hand side of Eq. (4.1.41), π^0 is an isospin state of $|1, 0\rangle$. Reaction of Eq. (4.1.41) can only happen through the isospin state $|1, 0\rangle$, because strong interaction does not change the isospin. So, it can be inferred that the amplitude of the reaction of Eq. (4.1.41) is only $1/\sqrt{2}$ times of Eq. (4.1.40). Therefore, the cross section of $\mathrm{p} + \mathrm{n} \to \pi^0 + \mathrm{d}$ is only half of that of $\mathrm{p} + \mathrm{p} \to \pi^+ + \mathrm{d}$. Such a prediction is consistent with experiment.

Therefore π^+, π^0, π^- belong to an isospin triplet state, of which the isospin is 1. The third isospin components of π^+, π^0, π^- are 1,0, and -1, respectively. The experimental measurement for the scattering processes in Eqs. (4.1.40) and (4.1.41) indicates that isospin conservation is consistent with truth.

In addition, for the isospin triplet composed by π mesons, the following relation is also satisfied:

$$Q = I_3 + B/2 \tag{4.1.43}$$

where the baryon number of pion is $B = 0$.

4.1.4.3. *The isospin of K mesosns*

There are four types of K mesons, which are K^+, K^-, K^0 and $\bar{\mathrm{K}}^0$. The mass of $\mathrm{K}^\pm$ is 493.7MeV, and the masses of K^0 and $\bar{\mathrm{K}}^0$ are 497.7MeV. The masses of them are very similar. Since K^0 and $\bar{\mathrm{K}}^0$ can be generated in pair in collision of nucleons, the strange number of them should be opposite. So K^0 and $\bar{\mathrm{K}}^0$ cannot be the same particle. The charges of K^0 and $\bar{\mathrm{K}}^0$ are the same, so it is not possible for them to be in the same isospin multiplet. For particles in the same isospin multiplet, their charges should differ by 1 successively. Therefore, the isospin of four kaons cannot be 3/2. In addition, because there are no more particles with similar masses to K mesons, the isospin of kaon cannot be larger than 3/2. So, it follows that the isospin of kaon should be $I < 3/2$, i.e., the possible value can be 0,1/2 or 1.

In addition, the following scattering processes can occur experimentally:

$$\pi^+ + \mathrm{n} \to \Lambda^0 + \mathrm{K}^+ \tag{4.1.44}$$

The mass of Λ^0 hyperon is 1115.7MeV. No other isotropic particles of Λ^0 with different charge, similar mass are found in experiment. So, the isospin of Λ^0 must be $I = 0$, and $I_3 = 0$. Then we compare the isospin of particles on both sides of the above equation

$$\begin{array}{lllll} & \pi^+ + & \mathrm{n} \to & \Lambda^0 + & \mathrm{K}^+ \\ I & 1 & 1/2 & 0 & ? \\ I_3 & 1 & -1/2 & 0 & ? \end{array} \tag{4.1.45}$$

The isospin of the left side is $I = 3/2$ or $1/2$, and right side $I = I_\mathrm{K}$, where I_K is the isospin of K^+. Since $I_\mathrm{K} \neq 3/2$, the isospin of K^+ can only be 1/2. So, the isospin of K^+ is $I = 1/2$, and $I_3 = 1/2$.

Similarly, through the scattering process

$$\mathrm{K}^- + \mathrm{p} \to \Lambda^0 + \pi^0 \tag{4.1.46}$$

the isospin of K^- can be determined to be $I = 1/2$, $I_3 = -1/2$.

From

$$\pi^- + \mathrm{p} \to \Lambda^0 + \mathrm{K}^0 \tag{4.1.47}$$

the isospin of K^0 can be obtained as $I = 1/2$, $I_3 = -1/2$. Since the third isospin component of antiparticle is opposite for sign, the isospin of $\bar{\mathrm{K}}^0$ is $I = 1/2$, and $I_3 = 1/2$.

The isospins of Kaons are listed in Table 4.1.

Table 4.1. The isospin and strange numbers of Kaons.

	K^+	K^0	$\bar{K}^0$	K^-
I	1/2	1/2	1/2	1/2
I_3	1/2	−1/2	1/2	−1/2
S	1	1	−1	−1

From Kaon's isospin $I = 1/2$ shown in Table 4.1, it can be known that four Kaons must be divided into two isospin doublets. K^0 and $\bar{K}^0$ can not be put into the same doublet, because the electric charges of them are the same. Then it can be inferred from the third isospin components and strange numbers that K^+ and K^0 make up one isospin doublet, and K^- and $\bar{K}^0$ make up another doublet.

Now, the charge Q, third component of isospin I_3 and baryon number B of each particle in such two doublets no longer satisfy the relation

$$Q = I_3 + B/2 \tag{4.1.48}$$

But if the strange number S is also considered, then Q, I_3, B and S can satisfy a similar relation

$$Q = I_3 + \frac{B+S}{2} \tag{4.1.49}$$

This generalized relation still holds for multiplet of p and n, and multiplet of pions, because the strange numbers of p, n and pions are $S = 0$. $B + S$ can be defined as a new quantum number, which is called hypercharge and denoted as Y

$$Y \equiv B + S \tag{4.1.50}$$

All the hypercharges of the particles in the same multiplet are the same, so Eq. (4.1.49) can be written as

$$Q = I_3 + Y/2 \tag{4.1.51}$$

This formula is called Gell-Mann–Nishijiama relation, which is an empirical relation about electric charge, the third component of isospin and hypercharge [2,3].

4.1.4.4. *Isospin multiplet of other particles*

The masses of Σ^+, Σ^0, Σ^- are similar. They constitute isospin triplet, the isospin of which are $I = 1$, $I_3 = 1, 0, -1$; hypercharge $Y = 0$, and strange number $S = -1$.

Ξ^0 and Ξ^- constitute isospin doublet, with isospin $I = 1/2$, $I_3 = 1/2$, $-1/2$; hypercharge $Y = -1$, strange number $S = -2$.

Δ^{++}, Δ^+, Δ^0, Δ^- constitute isospin quartet, with isospin $I = 3/2$, $I_3 = 3/2, 1/2, -1/2, -3/2$; hypercharge $Y = 1$, strange number $S = 0$.

4.1.5. *An example of isospin analysis of physical processes*

Let's take the πN scattering process as an example to see how the isospin analysis plays a role in predicting the relation between the cross sections of πN scattering.

There are six different physical states of πN system, which are π^+p, π^+n, π^0p, π^0n, π^-p and π^-n. These physical states can be coupled into isospin eigenstates. The results are

$$\begin{aligned}
|\pi^+\mathrm{p}\rangle &= \left|\frac{3}{2}, \frac{3}{2}\right\rangle \\
|\pi^+\mathrm{n}\rangle &= \sqrt{\frac{1}{3}}\left|\frac{3}{2}, \frac{1}{2}\right\rangle + \sqrt{\frac{2}{3}}\left|\frac{1}{2}, \frac{1}{2}\right\rangle \\
|\pi^0\mathrm{p}\rangle &= \sqrt{\frac{2}{3}}\left|\frac{3}{2}, \frac{1}{2}\right\rangle - \sqrt{\frac{1}{3}}\left|\frac{1}{2}, \frac{1}{2}\right\rangle \\
|\pi^0\mathrm{n}\rangle &= \sqrt{\frac{2}{3}}\left|\frac{3}{2}, -\frac{1}{2}\right\rangle + \sqrt{\frac{1}{3}}\left|\frac{1}{2}, -\frac{1}{2}\right\rangle \\
|\pi^-\mathrm{p}\rangle &= \sqrt{\frac{1}{3}}\left|\frac{3}{2}, -\frac{1}{2}\right\rangle - \sqrt{\frac{2}{3}}\left|\frac{1}{2}, -\frac{1}{2}\right\rangle \\
|\pi^-\mathrm{n}\rangle &= \left|\frac{3}{2}, -\frac{3}{2}\right\rangle
\end{aligned} \tag{4.1.52}$$

where the coefficients can be obtained by referring to the table of Clebsch–Gordan coefficients.

According to the invariance of strong interaction under isospin transformation, the Hamiltonian of a strong interacting system only depends on the total isospin of the system I, and there is no any dependence on the third component of isospin I_3. Therefore, for the πN system, there are only two basic scattering matrix elements, which are denoted as

$$\begin{aligned} A_{3/2} &= \langle I = 3/2 |\, \hat{H}\, | I = 3/2 \rangle \\ A_{1/2} &= \langle I = 1/2 |\, \hat{H}\, | I = 1/2 \rangle \end{aligned} \tag{4.1.53}$$

Then for all the possible πN scattering processes, the scattering amplitudes are

$$\begin{aligned} \langle \pi^+\mathrm{p}|\, \hat{H}\, |\pi^+\mathrm{p}\rangle &= A_{3/2} \\ \langle \pi^+\mathrm{n}|\, \hat{H}\, |\pi^+\mathrm{n}\rangle &= \frac{1}{3}A_{3/2} + \frac{2}{3}A_{1/2} \\ \langle \pi^+\mathrm{n}|\, \hat{H}\, |\pi^0\mathrm{p}\rangle &= \frac{\sqrt{2}}{3}A_{3/2} - \frac{\sqrt{2}}{3}A_{1/2} \\ \langle \pi^0\mathrm{p}|\, \hat{H}\, |\pi^0\mathrm{p}\rangle &= \frac{2}{3}A_{3/2} + \frac{1}{3}A_{1/2} \\ \langle \pi^0\mathrm{n}|\, \hat{H}\, |\pi^0\mathrm{n}\rangle &= \frac{2}{3}A_{3/2} + \frac{1}{3}A_{1/2} \\ \langle \pi^0\mathrm{n}|\, \hat{H}\, |\pi^-\mathrm{p}\rangle &= \frac{\sqrt{2}}{3}A_{3/2} - \frac{\sqrt{2}}{3}A_{1/2} \\ \langle \pi^-\mathrm{p}|\, \hat{H}\, |\pi^-\mathrm{p}\rangle &= \frac{1}{3}A_{3/2} + \frac{2}{3}A_{1/2} \\ \langle \pi^-\mathrm{n}|\, \hat{H}\, |\pi^-\mathrm{n}\rangle &= A_{3/2} \end{aligned} \tag{4.1.54}$$

For the inelastic scattering processes in the above amplitudes, such as $\pi^0\mathrm{p} \to \pi^+\mathrm{n}$ and $\pi^+\mathrm{n} \to \pi^0\mathrm{p}$, the magnitudes of the amplitudes are equal because of reciprocity. Therefore, for such kinds of processes, we need to only list one of them. From the above amplitudes, we can find the following relations:

$$\begin{aligned} \sigma_{\pi^+\mathrm{p} \to \pi^+\mathrm{p}} &= \sigma_{\pi^-\mathrm{n} \to \pi^-\mathrm{n}} \\ \sigma_{\pi^+\mathrm{n} \to \pi^+\mathrm{n}} &= \sigma_{\pi^-\mathrm{p} \to \pi^-\mathrm{p}} \\ \sigma_{\pi^0\mathrm{p} \to \pi^+\mathrm{n}} &= \sigma_{\pi^-\mathrm{p} \to \pi^0\mathrm{n}} \\ \sigma_{\pi^0\mathrm{p} \to \pi^0\mathrm{p}} &= \sigma_{\pi^0\mathrm{n} \to \pi^0\mathrm{n}} \end{aligned} \tag{4.1.55}$$

For different center-of-mass energy of πN system, the contributions of $A_{3/2}$ and $A_{1/2}$ are not of the same significance. If there exists an isospin state of

$I = 3/2\,(1/2)$ around a certain center-of-mass energy, then the amplitude with $I = 3/2\,(1/2)$ will dominate at this energy. There is the resonance $\Delta(1232)$ at the center-of-mass energy $E = 1240\text{MeV}$, of which the isospin is $I = 3/2$, spin $J = 3/2$ and this is the so-called resonance P_{33}, so the amplitude $A_{3/2}$ will be greatly larger than $A_{1/2}$. Therefore, we can ignore $A_{1/2}$ in Eq. (4.1.54), and the following relation can be obtained:

$$\sigma_{\pi^+\text{p} \to \pi^+\text{p}} : \sigma_{\pi^-\text{p} \to \pi^0\text{n}} : \sigma_{\pi^-\text{p} \to \pi^-\text{p}} = 1 : \frac{2}{9} : \frac{1}{9} \tag{4.1.56}$$

It is experimentally found that, around this resonance, the ratio of the scattering cross sections of π^+p and π^-p reactions is 3:1, which is consistent with the relation in Eq. (4.1.56).

The importance of the concept of isospin lies in the discovery of a symmetry of strong interactions — isospin symmetry. With this symmetry, even if we don't know the exact detail of the strong interaction, we can still predict the relation between some physical observables.

4.2. Exchange Symmetry: Generalized Identity Principle

The wave function describing a completely identical multiparticle system is either symmetric or antisymmetric under any two-particle exchanges. The exchange symmetry of wave functions does not change with time. Particles that need to be described in terms of symmetric wave functions are bosons, while particles that need to be described by antisymmetric wave functions are fermions.

Without considering the spin–orbit coupling, the wave function of an identical multiparticle system is the product of the space wave function and the spin wave function

$$\Phi\left(r_1, s_1, r_2, s_2, \ldots, r_n, s_n\right) = \phi\left(r_1, r_2, \ldots, r_n\right) \chi_s\left(s_1, s_2, \ldots, s_n\right) \tag{4.2.1}$$

where ϕ is the space wave function and χ_s the spin wave function.

Considering the isospin symmetry of strong interaction, the above symmetry can be extended. Regarding particles in the same isospin multiplet as identical particles in a generalized sense, and multiplying the isospin wave function to the space and spin wave functions, we can generalize the total wave function as

$$\Phi\left(r, s, I\right) = \phi(r)\chi_s(s)\chi_I(I) \tag{4.2.2}$$

where r, s, I denote $(r_1, r_2, \ldots, r_n)$, $(s_1, s_2, \ldots, s_n)$ and $(I_1, I_2, \ldots, I_n)$, respectively, and $\chi_I(I)$ is the isospin wave function.

For particles in the same isospin multiplet, the extended wave function is antisymmetric for any two-fermion exchange, while it is symmetric for any two-boson exchange. This is called the generalized identical principle.

Explanations: (1) For true identical particles, because the isospin wave functions of identical particles are always symmetric, the generalized identical principle does not add anything new. It is completely equivalent to the identical principle for identical particles.

(2) For particles within an isospin multiplet that are not really identical, such as π^+, π^- and π^0, the possible isospin states of this particle system shall be restricted. For example, for π or ρ meson system, a symmetry factor $(-1)^L$ shall emerge from the space wave function, and factors $(-1)^S$ and $(-1)^I$ from spin and isospin wave functions, respectively, under transformation of particle exchange. So, the total exchange factor is $(-1)^{L+S+I}$. Because mesons π and ρ are bosons, the total wave function should be symmetric under two-particle exchange transformation. Then we have $L + S + I = 2n$, where n is a non-negative integer; similarly, for a system composed of two nucleons, we must have $L + S + I = 2n+1$.

As an example of the application of the generalized identical principle and isospin conservation, let's analyze the decays of $\rho \to \pi\pi$ and $\omega \to \pi\pi$. Both of the two decays proceed via strong interaction. The masses of ρ and ω are very similar, $m_\rho = (769.3 \pm 0.8)\text{MeV}$ and $m_\omega = (782.57 \pm 0.12)\text{MeV}$. But ρ meson decays to $\pi\pi$ with decay rate almost of 100%, while the branching ratio of decay $\omega \to \pi\pi$ is only 2.2%. Here, the branching ratio is defined as the ratio of partial decay width of a specific decay mode to the total decay width of the decaying particle. So, the branching ratio stands for the decay rate of a particle decaying into specific final states.

There are three types of ρ mesons, which are ρ^+, ρ^0 and ρ^-. They constitute an isospin triplet. The isospin of ρ meson is $I_\rho = 1$. The third isospin components of ρ^+, ρ^0 and ρ^- are 1, 0 and -1, respectively, while ω meson is an isospin singlet, the isospin of which is 0. ρ and ω are all vectors with spin $J = 1$.

Because all the spins of ρ and ω are 1, according to angular momentum conservation, the relevant orbital angular momentum of $\pi\pi$ must be $L = 1$ (note the spin of π meson is 0). Therefore, the symmetric factor of the space wave function of $\pi\pi$ is $(-1)^L = -1$, i.e., the space wave function of $\pi\pi$ is antisymmetric. Since π meson is boson, the total wave function of which must be symmetric, the isospin wave function of $\pi\pi$ must also be antisymmetric. However, with the isospin wave function being antisymmetric, the total isospin of $\pi\pi$ system can only be 1. Therefore, the decay of ρ to $\pi\pi$ conforms to isospin conservation, then the decay branching ratio can be large. However, ω decaying to $\pi\pi$ violates isospin conservation, which can only be an isospin violating effect. So, the branching ratio of this decay mode must be tiny, which is consistent with the experimental measurement. This example also shows that isospin symmetry is an approximate symmetry, which can be damaged by a few percent.

4.3. Isospin Violation

Electric charge, baryon number and strange number are all conserved in electromagnetic interactions. So hypercharge should also be conserved. Then according to Gell-Mann–Nishijima relation

$$Q = I_3 + Y/2 \tag{4.3.1}$$

the third component of isospin I_3 is also conserved in electromagnetic interaction. A large number of experiments have shown that the total isospin I is not conserved in electromagnetic interaction, but the change can only be 0 or 1.

Therefore, the selection rule of isospin change under electromagnetic interaction can be summarized as follows:

$$|\Delta I_3| = 0, \quad |\Delta I| = 0 \quad \text{or} \quad 1 \tag{4.3.2}$$

According to the selection rule above, we examine two electromagnetic decay processes: $\pi^0 \to \gamma\gamma$ and $\Sigma^0 \to \Lambda + \gamma$.

(1) The isospin of the initial and final states of the process $\pi^0 \to \gamma\gamma$ is labeled in the following:

$$\begin{array}{lllll} & \pi^0 \to & & \gamma & \gamma \\ I & 1 & & 0 & 0 \\ I_3 & 0 & & 0 & 0 \end{array}$$

so, in this process $|\Delta I| = 1$, $|\Delta I_3| = 0$.

(2) The isospin of the process $\Sigma^0 \to \Lambda + \gamma$ is labeled as

	$\Sigma^0 \to$	$\Lambda +$	γ
I	1	0	0
I_3	0	0	0

The selection rule $|\Delta I| = 1$, $|\Delta I_3| = 0$ is also satisfied.

Both the total isospin and the third component are not conserved in weak interaction, but a large number of experiments have shown that $|\Delta I| \leq 1$ always holds in weak interaction. Therefore, we can hypothesize that this constraint always holds in weak interaction. This is called the "minimal violation" hypothesis. In general, the selection rule $|\Delta I| = 1/2$ always holds in weak interaction [4].

4.4. *G* Parity

As is known that strong interaction is invariant under C and isospin transformation, but under charge conjugation, the nucleon system transforms as

$$\hat{C}\begin{pmatrix} \mathrm{p} \\ \mathrm{n} \end{pmatrix} = \begin{pmatrix} \bar{\mathrm{p}} \\ \bar{\mathrm{n}} \end{pmatrix} \tag{4.4.1}$$

If the isospin state of proton p is denoted as ↑, the isospin state of $\bar{\mathrm{p}}$ can be denoted as ↓. Therefore, Eq. (4.4.1) indicates that the isospin state of the nucleon system is changed after C transformation. Because isospin transformation is symmetric operation of strong interaction, and the isospin state of the system can be rotated in the isospin space in isospin transformation, it is possible that the isospin state can be recovered back to its original state by choosing certain isospin transformation angle $\vec{\theta}$, without affecting the strong interaction of the system. Such a transformation is possibly a symmetric transformation for the system. Therefore, a new quantum number can be introduced in this way, which is also conserved in strong interaction. This new quantum number, set up in the joint transformation of isospin and charge conjugation, is called G parity [5, 6].

We shall prove that the operator for this G transformation is

$$\hat{G} = \hat{C} e^{\mathrm{i}\pi \cdot \hat{I}_2} \tag{4.4.2}$$

where $\hat{I}_2$ is the isospin rotation operator around the second axis, and $\hat{C}$ the charge conjugation operator. Note that all the operators here are in Hilbert space. Equation (4.4.2) shows that G transformation is a joint operation of rotating 180° around the second isospin axis and the C transformation. What needs to be proved is that, after G transformation, the isospin state of the system is indeed recovered.

Corresponding to the isospin transformation operator in Hilbert space

$$\hat{S} = e^{\mathrm{i}\hat{\vec{I}}\cdot\vec{\theta}} \tag{4.4.3}$$

where $\hat{\vec{I}} \cdot \vec{\theta} \equiv \left(\hat{I}_1\theta_1 + \hat{I}_2\theta_2 + \hat{I}_3\theta_3\right)$, with $\hat{\vec{I}}$ being the isospin operator in Hilbert space, there exists the relevant isospin transforming matrix $\vec{\tau}/2$. Then the total isospin transforming matrix is

$$S = e^{\mathrm{i}\frac{1}{2}\vec{\tau}\cdot\vec{\theta}} \tag{4.4.4}$$

where $\vec{\tau} = (\tau_1, \tau_2, \tau_3)$ is Hermitian matrix. The collection of S matrices makes up the SU(2) group. For the case of dimension 2, τ_i's are Pauli matrices

$$\tau_1 = \sigma_1, \quad \tau_2 = \sigma_2, \quad \tau_3 = \sigma_3 \tag{4.4.5}$$

For a field operator $\psi(x)$ in Hilbert space, its isospin transformation is

$$e^{\mathrm{i}\hat{\vec{I}}\cdot\vec{\theta}}\psi(x)e^{-\mathrm{i}\hat{\vec{I}}\cdot\vec{\theta}} = e^{\mathrm{i}\frac{1}{2}\vec{\tau}\cdot\vec{\theta}}\psi(x) \tag{4.4.6}$$

Now, let us consider dimension-2 nucleon system $\begin{pmatrix} |p\rangle \\ |n\rangle \end{pmatrix}$. Acting $\hat{G}$ on it

$$\hat{G}\begin{pmatrix} |\mathrm{p}\rangle \\ |\mathrm{n}\rangle \end{pmatrix} = \hat{C}e^{\mathrm{i}\pi\frac{\sigma_2}{2}}\begin{pmatrix} |p\rangle \\ |n\rangle \end{pmatrix} \tag{4.4.7}$$

According to Eq. (4.1.29), we have

$$e^{\mathrm{i}\pi\frac{\sigma_2}{2}} = \mathrm{i}\sigma_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{4.4.8}$$

Substituting Eqs. (4.4.8) into (4.4.7), we get

$$\hat{G}\begin{pmatrix} |\mathrm{p}\rangle \\ |\mathrm{n}\rangle \end{pmatrix} = \begin{pmatrix} |\bar{\mathrm{n}}\rangle \\ -|\bar{\mathrm{p}}\rangle \end{pmatrix} \tag{4.4.9}$$

Equation (4.4.9) indicates that the nucleon system indeed returns back to its original isospin state after G transformation.

Let's consider pion system further. There are three types of electric charge states for pion, π^+, π^0 and π^-, which can be arranged in the form of a column matrix

$$\begin{pmatrix}\pi^+\\ \pi^0\\ \pi^-\end{pmatrix} \tag{4.4.10}$$

To consider the isospin transformation of the pion system, we need a three-dimensional representation of the SU(2) group. From the theory of group representation, we know that, in three-dimensional representation, the three generators of the SU(2) group are

$$\frac{1}{2}\tau_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}0 & -1 & 0\\ -1 & 0 & 1\\ 0 & 1 & 0\end{pmatrix},\ \frac{1}{2}\tau_2 = \frac{\mathrm{i}}{\sqrt{2}}\begin{pmatrix}0 & 1 & 0\\ -1 & 0 & -1\\ 0 & 1 & 0\end{pmatrix},\ \frac{1}{2}\tau_3 = \begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & -1\end{pmatrix} \tag{4.4.11}$$

So, for the three-dimensional system, $\hat{G}$ should be

$$\hat{G} = \hat{C}e^{\mathrm{i}\pi\frac{\tau_2}{2}} \tag{4.4.12}$$

We need to calculate $e^{\mathrm{i}\pi\frac{\tau_2}{2}}$ at first. Using

$$e^{\mathrm{i}\pi\frac{\tau_2}{2}} = \sum_{n=0}^{\infty}\frac{1}{n!}\left(i\pi\frac{\tau_2}{2}\right)^n \tag{4.4.13}$$

and the relations (which can be verified by direct calculation and recursion)

$$\left(i\frac{\tau_2}{2}\right)^{2m+1} = (-1)^m\mathrm{i}\frac{\tau_2}{2},\ \text{integer of}\quad m \geq 0 \tag{4.4.14}$$

$$\left(i\frac{\tau_2}{2}\right)^{2m} = (-1)^{m+1}\left(\mathrm{i}\frac{\tau_2}{2}\right)^2,\ \text{integer of}\quad m \geq 1 \tag{4.4.15}$$

we can finally obtain

$$\begin{aligned} e^{\mathrm{i}\pi\frac{\tau_2}{2}} &= \mathrm{i}\frac{\tau_2}{2}\sin\pi + \left(\mathrm{i}\frac{\tau_2}{2}\right)^2(1-\cos\pi) + 1\\ &= \begin{pmatrix}0 & 0 & -1\\ 0 & -1 & 0\\ -1 & 0 & 0\end{pmatrix} \end{aligned} \tag{4.4.16}$$

Table 4.2. G parities of some mesons.

	π	ρ	ω	ϕ	J/ψ
G	-1	$+1$	-1	-1	-1

Therefore, for the pion system

$$\hat{G}\begin{pmatrix}\pi^+\\ \pi^0\\ \pi^-\end{pmatrix} = \hat{C}\begin{pmatrix}0 & 0 & -1\\ 0 & -1 & 0\\ -1 & 0 & 0\end{pmatrix}\begin{pmatrix}\pi^+\\ \pi^0\\ \pi^-\end{pmatrix}$$

$$= -\begin{pmatrix}\pi^+\\ \pi^0\\ \pi^-\end{pmatrix} \tag{4.4.17}$$

The above shows that the pion system recovers its original isospin state, and in addition, the eigenvalue of G parity of pion is -1. Then we can say that the G parity of pion is -1. $\hat{C}\,|\pi^\pm\rangle = |\pi^\mp\rangle$ has been used in Eq. (4.4.17), which indicates that the phase angle of the charged pion under C transformation is taken to be the same as that of π^0 under charge conjugation. The phase angle of π^0 under charge conjugation is an absolute angle. The charge parity of π^0 is $+1$.

Because C parity and isospin are all conserved in strong interaction, G parity is also conserved. According to conservation of G parity, we can determine G parity of ρ meson using the processes $\rho^+ \to \pi^+\pi^0$, $\rho^0 \to \pi^+\pi^-$, and $\rho^- \to \pi^-\pi^0$.

Because G parity is a multiplicative quantum number, and $\rho \to \pi\pi$ is a strong decay process, we can determine that G parities of ρ^+, ρ^0 and ρ^- are all 1. Therefore, the following selection rule can be inferred:

$$\rho \not\to 3\pi \tag{4.4.18}$$

G parities of some other particles are given in Table 4.2.

For a particle constituted of fermion and anti-fermion and with integer isospin, its G parity is generally $G = (-1)^{L+S+I}$, where L is the relative orbital angular momentum of the fermion and anti-fermion, S the total spin of the fermion pair and I the isospin of this particle.

References

1. Heisenberg W. *Zeit. Phys.*, 1932, 77: 1.
2. Nishijima K. *Prog. Theor. Phys.*, 1955, 13: 285.
3. Gell-Mann M. *Nuovo Cimento*, 1956, 4: 848.
4. Gell-Mann M, Pais A. *Proceeding of the Glasgow Conference on Nuclear and Meson Physics*, 1954: 342.
5. Lee T D, Yang C N. *Nuovo Cimento*, 1956, 3: 749.
6. Goebel C. *Phys. Rev.*, 1956, 103: 258.

This page intentionally left blank

Chapter 5

Quark Model of Hadrons

Until now, hadrons (including resonances) discovered are about one hundred. Most of them have very short lifetimes, but the differences of their lifetimes are very large, e.g.,

Δ (1232)	$\tau \sim 10^{-23}$ s
p	$\tau > 10^{32}$ yr (probably infinity)
n	$\tau \sim 15$ min
π^0	$\tau \sim 0.8 \times 10^{-16}$ s
$\pi^\pm$	$\tau \sim 2.6 \times 10^{-8}$ s

It is found experimentally that these hadrons can be divided into different groups according to their isospin I and strangeness S. The masses of the hadrons within the same isospin multiplet are similar. We give some examples as follows:

π^- π^0 π^+	$I = 1$	$S = 0$	$J^P = 0^-$
K^0 K^+	$I = 1/2$	$S = 1$	$J^P = 0^-$
K^- $\bar{\mathrm{K}}^0$	$I = 1/2$	$S = -1$	$J^P = 0^-$
ρ^- ρ^0 ρ^+	$I = 1$	$S = 0$	$J^P = 1^-$
K^{*0} K^{*+}	$I = 1/2$	$S = 1$	$J^P = 1^-$
K^{*-} $\bar{\mathrm{K}}^{*0}$	$I = 1/2$	$S = -1$	$J^P = 1^-$
p n	$I = 1/2$	$S = 0$	$J^P = \left(\frac{1}{2}\right)^+$
Σ^- Σ^0 Σ^+	$I = 1$	$S = -1$	$J^P = \left(\frac{1}{2}\right)^+$
Ξ^- Ξ^0	$I = 1/2$	$S = -2$	$J^P = \left(\frac{1}{2}\right)^+$
Δ^- Δ^0 Δ^+ Δ^{++}	$I = 3/2$	$S = 0$	$J^P = \left(\frac{3}{2}\right)^+$

The discovery of large number of hadrons naturally reminds us of the periodic table of the elements. Are these hadrons all “elementary”? Are they possibly made of a small number of more fundamental objects?

In 1949, Fermi and Yang proposed the Fermi-Yang model [1], They assumed that p and n are fundamental, π mesons were consisted of p, n and their antiparticles, that is, $\pi^+ = \mathrm{p\bar{n}}$, $\pi^- = \mathrm{\bar{p}n}$, $\pi^0 = \frac{1}{2}(\mathrm{p\bar{p}} - \mathrm{n\bar{n}})$.

In 1956, Sakata proposed Sakata model [2], it assumed that only the three particles p, n, Λ were fundamental, and these three objects constituted the fundamental representation of the SU(3) symmetry. Sakata model explains the structures of π and K mesons successfully, but it encounters serious difficulty in explaining the structure of baryons. Because the baryon number of p, n, Λ $B = 1$, in order to construct baryons ($B = 1$) we need to use two of p, n, Λ and one of their antiparticles, that is, from group representation point of view, we need to use $3 \otimes 3 \otimes 3^*$ for constructing baryons. In this way, there will be many baryons which do not exist in nature. So Sakata model is ruled out.

In 1961, Gell-Mann and Ne’eman abandoned Sakata model and directly considered that hadrons themselves possibly have SU(3) symmetry. They put p, n, Λ, Σ, Ξ into octet of SU(3) group, that is so-called eightfold way [3]. They also put Δ, Σ^*, Ξ^*, etc. into SU(3) decuplet, and predicted the existence of Ω^- which is missing in the decuplet filling. They further predicted that the mass of the Ω^- is 1672 MeV. In 1964, Ω^- was observed at Brookhaven national lab [4], the measured mass is consistent with prediction. This means that SU(3) symmetry is correct. In the same year, Gell-Mann and Zweig proposed the quark model [5]. In this model, the three quarks u, d, s with different flavor fill into the SU(3) fundamental representation (denoted by 3), while mesons and baryons fill into the octet of $3 \otimes 3^* = 8 + 1$ and the octet, decuplet of the representation $3 \otimes 3 \otimes 3$.

SU(3) symmetry gets great success and is proved to be the correct model for hadron structure. In this chapter, we shall discuss the quark model in detail. For this purpose, we need to introduce some mathematics first.

5.1. Mathematical Basics

5.1.1. *Decomposition of SU(n) group representation product, Young tableau*

SU(n) group is the group of $n \times n$ unitary matrices with unit determinant. If we use an $n \times n$ matrix u to denote an arbitrary element of the group, then

$$uu^+ = 1, \quad \det u = 1 \tag{5.1.1}$$

an $n \times n$ complex matrix has $2n^2$ real parameters. $uu^+ = 1$ gives $n + \frac{n^2-n}{2} \times 2 = n^2$ constrains, $\det u = 1$ also gives one constrain, so for $2n^2$ real parameters, only $n^2 - 1$ real parameters are independent; therefore SU(3) group has $n^2 - 1$ generators. So the element u can be expressed by

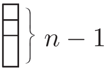
$$u = \mathrm{e}^{\mathrm{i}T_i\theta_i}, \quad i = 1, 2, \ldots, n^2 - 1 \tag{5.1.2}$$

where θ_i are $n^2 - 1$ independent real parameters, T^i are $n^2 - 1$ generators whose concrete forms will be given later. According to Eq. (5.1.1), T_i must be Hermitian and traceless

$$T_i^+ = T_i, \quad \mathrm{Tr}\ T_i = 0 \tag{5.1.3}$$

Now we shall explain how to use the Young tableau for the decomposition of the SU(n) group representation product, the proof can be found in [6].

The dimension of irreducible representation: The fundamental representation of the SU(n)group is denoted by a small box □, i.e. the representation $\underline{n}$, while $n - 1$ box column



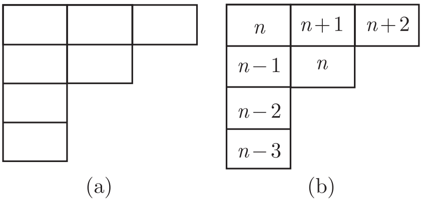
is its conjugate representation $\underline{n}^*$.

Any irreducible representation of the SU(n)group can be denoted by a Young tableau. For example, in Fig. 5.1, (a) denotes a irreducible representation. In order to calculate its dimension, we first fill number n in each box along the diagonal line of the tableau (a), the numbers filled in the boxes of each row need to increase one unit from left to right, the numbers filled in each boxes of a column need to decrease one unit from top to bottom, this is shown in (b) of Fig. 5.1.

The product of the numbers in all boxes in Fig. 5.1(b) is denoted by $c(n)$:

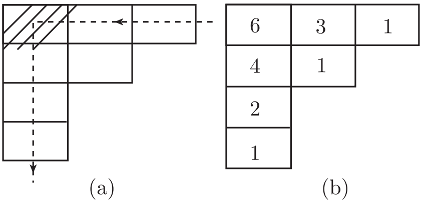
$$c(n) = n^2(n + 1)(n + 2)(n - 1)(n - 2)(n - 3) \tag{5.1.4}$$

Now we consider another set of numbers. For each box in Fig. 5.1(a), for instance the box with oblique line, draw two perpendicular lines, in the shape of a "hook", one going to the right and another going downward as



(a) (b)

Fig. 5.1. Calculation for the dimension of irreducible representation.



(a) (b)

Fig. 5.2. Hooke length.

shown in Fig. 5.2(a). The total number of the boxes penetrated by this hook, including the originating box itself, is called the "hook length" of the box with oblique line. Every box in Fig. 5.2(a) has its oven hook length, we label them in Fig. 5.2(b). The product of all the hook lengths is denoted by $H(n)$,

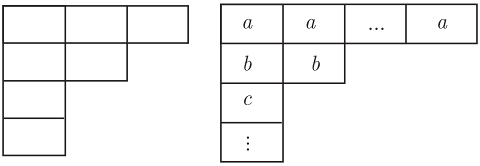
$$H(n) = 6 \cdot 4 \cdot 3 \cdot 2 \cdot 1 \cdot 1 \cdot 1 \tag{5.1.5}$$

then the dimension of the Young tableau (a) is

$$D = \frac{c(n)}{H(n)} = \frac{n^2(n+1)(n+2)(n-1)(n-2)(n-3)}{6 \times 4 \times 3 \times 2 \times 1 \times 1 \times 1} \tag{5.1.6}$$

Decomposition for the product of two representations: In Fig. 5.3 we present two Young tableaux (a) and (b) for two representations. Now we decompose the product of the two representations. The procedure is as follows:

(1) In the Young tableau (b) on the right-hand side, we label the symbol "a" in each box of the first row, "b" in each box of the second row, "c" in each box of the third row and so on.
(2) Add the boxes labeled "a" in Fig. 5.3(b) into the Young tableau (a) in all possible ways successively, but each adding must gets a meaningful



(a) (b)

Fig. 5.3. The decomposition for the product of two representations.

Young tableau (i.e. the length of rows does not increase going from top to bottom, etc.). After that repeat the same process with the symbol "b"s, "c"s, ...

(3) In the second step, boxes with the same symbol are not allowed to appear in same column.

(4) In the above steps, the Young tableau with more than n rows (for SU(n) group) must be abandoned.

(5) After finishing all the above steps we only keep the Young tableaux satisfying the following conditions: if we count the symbols in the boxes from right to left and from top to bottom, the total number of symbols "a"s must be large or equal to that of "b"'s, the total number of symbols "b"'s must be large or equal to that of "c"'s, etc.

(6) We only keep the qualified Young tableaux in which from left to right the total number of boxes of the first column $\geq$ that of the second column $\geq$ that of the third column $\geq \cdots$, and abandon the columns in which the number of rows exceeds n.

For the Young tableaux obtained in the last step, if two tableaux have the same pattern and labeled symbols, for instance, symbols "a", "b", etc., they are considered as the same tableau, the two tableaux with the same pattern but different labeled symbols will be considered as different tableaux.

For example, consider the product of the two Young tableaux $\square$ and $\square\square$.

(1) $\square \otimes \boxed{a}\boxed{a}$

(2) add one $\boxed{a}$ into the tableau on the left and obtain

$\square\boxed{a}$ and $\begin{array}{|c|}\hline \ \\ \hline a \\ \hline \end{array}$,

then add second $\boxed{a}$,

$\begin{array}{|c|c|}\hline \ & a \\ \hline\end{array}$ after adding the second $\boxed{a}$ becomes $\begin{array}{|c|c|c|}\hline \ & a & a \\ \hline\end{array} + \begin{array}{|c|c|}\hline \ & a \\ \hline a & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}$,

$\begin{array}{|c|}\hline \ \\ \hline a \\ \hline\end{array}$ after adding the second $\boxed{a}$ becomes $\begin{array}{|c|c|}\hline \ & a \\ \hline a & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}$, then according to the step (3) above

$\begin{array}{|c|}\hline \ \\ \hline a \\ \hline a \\ \hline\end{array}$ should be abandoned, according to step (5)

$$\begin{array}{|c|}\hline \ \\ \hline\end{array} \otimes \begin{array}{|c|c|}\hline \ & \ \\ \hline\end{array} = \begin{array}{|c|c|c|}\hline \ & \ & \ \\ \hline\end{array} + \begin{array}{|c|c|}\hline \ & \ \\ \hline \ & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} \tag{5.1.7}$$

this is because that the two $\begin{array}{|c|c|}\hline \ & a \\ \hline a & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}$ obtained in (2) have the same pattern and the same labeled symbols are considered as the same tableau, therefore in Eq. (5.1.7) there is no the factor "2" in front of $\begin{array}{|c|c|}\hline \ & \ \\ \hline \ & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}$.

Previously, we only discussed the calculation of representation dimension and the decomposition for the product of two representations in SU(n) scheme. We leave the decomposition of SU(n) representation into representation of SU(m)× SU($n-m$) to the sections of grand unified theories later.

5.1.2. *Tensor analysis of the SU(3) group*

We skip the discussion for the relation between tensors and the Young tableaux, we shall go to SU(3) group tensor analysis directly for the use in SU(3) quark model.

Define $T^{i_1 i_2 \cdots i_n}_{a_1 a_2 \cdots a_m}$ as the component of a rank (n, m) tensor. The upper indices of it are called covariant indices and transform according to SU(3) group element u. While the lower indices are called contravariant indices and transform according to u^+. So under SU(3) transformation

$$T^{i_1 i_2 \cdots i_n}_{a_1 a_2 \cdots a_m} \to u^{i_1}_{j_1} u^{i_2}_{j_2} \cdots u^{i_n}_{j_n} \left(u^+\right)^{b_1}_{a_1} \left(u^+\right)^{b_2}_{a_2} \cdots \left(u^+\right)^{b_m}_{a_m} T^{j_1 j_2 \cdots j_n}_{b_1 b_2 \cdots b_m} \tag{5.1.8}$$

Tensors which have only upper indices are called covariant tensors, tensors which have only lower indices are called contravariant tensors, while tensors which have both upper and lower indices are named mixed tensors. The tensor in Eq. (5.1.8) is just mixed one of rank (n, m).

(1) The addition or subtraction of two tensors with the same rank still get the same rank tensor.

(2) The product of a rank (n, m) tensor and a rank (q, p) tensor will get a new tensor of rank $(n+p, m+q)$, i.e. the produced new tensor is

$$A^{i_1 i_2 \cdots i_n}_{a_1 a_2 \cdots a_m} \cdot B^{j_1 j_2 \cdots j_q}_{b_1 b_2 \cdots b_p}$$

(3) Isotropic tensors δ^i_a (also called Kronecker delta), $\varepsilon^{ijk}, \varepsilon_{ijk}$ (also called Levi-Civita symbol).

It is easy to prove that the mixed tensor of rank $(1,1)$ δ^i_a, the covariant tensor of rank (3, 0) ε^{ijk}, and the contravariant tensor of rank (0, 3) ε_{ijk} are all invariant under SU(3) transformation.

According to definition

$$\delta^i_a = \begin{cases} 1, \ i = a \\ 0, \ i \neq a \end{cases} \tag{5.1.9}$$

$$\varepsilon^{ijk} = \varepsilon_{ijk} = \begin{cases} 1, & ijk \ \text{is} \ 123 \ \text{even permutation} \\ -1 & ijk \ \text{is} \ 123 \ \text{odd permutation} \\ 0, & \text{otherwise} \end{cases} \tag{5.1.10}$$

Now we prove that Eqs. (5.1.9) and (5.1.10) are invariant under SU(3) transformation.

Under SU(3) transformation

$$\begin{aligned} \delta^i_a &\to u^i_j \left(u^+\right)^b_a \delta^j_b = (uu^+)^i_a = \delta^i_a \\ \varepsilon^{ijk} &\to u^i_a u^j_b u^k_c \varepsilon^{abc} = \varepsilon^{ijk} \end{aligned}$$

In the last step $\det u = 1$ is used, and the exchange for any two indices of ijk is equivalent to the two row exchange in $\det u$, so the sign is changed. Similarly, we can prove that ε_{ijk} is also SU(3) invariant.

(4) *Tensor contraction*: The indices of a tensor $T^{i_1 i_2 \cdots i_n}_{a_1 a_2 \cdots a_m}$ of rank (n, m) can be contracted by use of $\delta^i_a, \varepsilon^{ijk}, \varepsilon ijk$, for example

$$\begin{aligned} &T^{i_1 \cdots i_a \cdots i_n}_{j_1 \cdots j_b \cdots j_m} \delta^{j_b}_{i_a} \to \text{a rank}(n-1, m-1) \text{ tensor} \\ &T^{i_1 \cdots i_a \cdots i_{a'} \cdots i_n}_{j_1 j_2 \cdots j_m} \varepsilon_{j_{m+1} i_a i_{a'}} \to \text{a rank } (n-2, m+1) \text{ tensor} \\ &T^{i_1 \cdots i_n}_{j_1 \cdots j_b \cdots j_{b'} j_m} \varepsilon^{i_{n+1} j_b j_{b'}} \to \text{a rank } (n+1, m-2) \text{ tensor} \end{aligned}$$

(5) *Irreducible and reducible tensor*: A tensor T of rank (n, m) is called reducible, whenever we can get a non-zero and lower rank (n', m') tensor form by the above contracting procedure, and $n' + m' < n + m$.

Obviously, if the tensor $T^{i_1 i_2 \cdots i_n}_{j_1 j_2 \cdots j_m}$ is irreducible, its upper indices must be totally symmetric, and its lower indices also must be totally symmetric, and furthermore the contraction of a upper and a lower index must be zero, that is the traceless condition,

$$T^{i_1 \cdots i_n}_{j_1 \cdots j_m} \delta^{j_1}_{i_1} = 0 \tag{5.1.11}$$

Note that not all the components of an irreducible tensors are necessarily independent.

(6) *Representation*: If the number of linearly independent components for a rank (n, m) tensor $T^{i_1 \cdots i_n}_{j_1 \cdots j_m}$ is d, then these d independent components span a d-dimensional tensor space, and d linearly independent tensors can be the basis of the SU(3) group. Assume that $T_1, T_2, \ldots, T_d$ are independent tensors, then any tensor T in this d-dimensional space can be written as

$$T = c_1 T_1 + c_2 T_2 + \cdots + c_d T_d = \sum_{i=1}^{d} c_i T_i \tag{5.1.12}$$

and T can be expressed as a column vector $\begin{pmatrix} c_1 \\ \vdots \\ c_d \end{pmatrix}$. Under SU(3) transformation, $T \to T'$ can be written as

$$\begin{pmatrix} c'_1 \\ c'_2 \\ \vdots \\ c'_d \end{pmatrix} = V \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_d \end{pmatrix} \tag{5.1.13}$$

The set of the transformation V $\{V\}$ form a d-dimensional representation of the SU(3) group. The product of the two tensors with dimensions d_1, d_2 respectively in the tensor space will form a new tensor; the new tensor spans a space with dimension

$$d = d_1 \times d_2 \tag{5.1.14}$$

This product is called direct product of the two tensor spaces. The new tensor in the direct product space is not always irreducible. It can be reduced into a direct sum of several groups of irreducible tensors. These groups of irreducible tensors span corresponding subspaces. The direct product space is reduced into direct sum of these several irreducible subspaces. Its corresponding representation is also reduced into the direct sum of several irreducible representations. The reduction of the direct product of two representations can be done either by use of Young tableaux or tensor analysis. We shall discuss it in detail later.

5.1.3. *SU(3) group generators and Casmir operators*

The set of the unitary unimodular 3×3 matrices $\{u\}$ forms SU(3) group, where

$$u^+u = 1, \quad \det u = 1, \quad u = \mathrm{e}^{\mathrm{i}\theta_i T_i} \quad (i = 1, 2, \ldots, 8) \tag{5.1.15}$$

For the 3×3 complex matrix representation, $T_i = \frac{1}{2}\lambda_i (i = 1, \ldots, 8)$ are the group generators, called Gell-Mann matrices. Their concrete expression is as follows:

$$\begin{gathered}
\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_2 = \begin{pmatrix} 0 & -\mathrm{i} & 0 \\ \mathrm{i} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\
\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \lambda_5 = \begin{pmatrix} 0 & 0 & -\mathrm{i} \\ 0 & 0 & 0 \\ \mathrm{i} & 0 & 0 \end{pmatrix}, \quad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \\
\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\mathrm{i} \\ 0 & \mathrm{i} & 0 \end{pmatrix}, \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}
\end{gathered} \tag{5.1.16}$$

Obviously, λ_i satisfies

$$\begin{gathered}
\lambda_i^+ = \lambda_i, \quad \mathrm{Tr}(\lambda_i \lambda_j) = 2\delta_{ij} \\
[\lambda_i, \lambda_j] = 2\mathrm{i} f_{ijk} \lambda_k, \quad [T_i, T_j] = \mathrm{i} f_{ijk} T_k \\
\{\lambda_i, \lambda_j\} = \frac{4}{3}\delta_{ij} + 2d_{ijk}\lambda_k
\end{gathered} \tag{5.1.17}$$

where the structure constants f_{ijk} are totally antisymmetric with respect to i, j, k, while d_{ijk} are totally symmetric with respect to i, j, k. The non-zero values of f_{ijk}, d_{ijk} are listed in Table 5.1.

SU(3) is a group of rank two, it has two Casmir operators C_2, C_3, and

$$[C_2, T_i] = 0, \quad [C_3, T_i] = 0, \quad i = 1, 2, \ldots, 8 \tag{5.1.18}$$

$$C_2 = \sum_{i=1}^{8} T_i T_i = \frac{1}{2}\{I_+, I_-\} + I_3^2 + \frac{1}{2}\{U_+, U_-\} + \frac{1}{2}\{V_+, V_-\} + T_8^2 \tag{5.1.19}$$

Table 5.1. The non-zero values of SU(3) group structure constants.

f_{123}	f_{147}	f_{156}	f_{246}	f_{257}	f_{345}	f_{367}	f_{458}	f_{678}
1	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{3}}{2}$
d_{118}	d_{146}	d_{157}	d_{228}	d_{247}	d_{256}	d_{338}	d_{344}	d_{355}
$\frac{1}{\sqrt{3}}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{\sqrt{3}}$	$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{\sqrt{3}}$	$\frac{1}{2}$	$\frac{1}{2}$
d_{366}	d_{377}	d_{448}	d_{558}	d_{668}	d_{778}	d_{888}		
$-\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2\sqrt{3}}$	$-\frac{1}{2\sqrt{3}}$	$-\frac{1}{2\sqrt{3}}$	$-\frac{1}{2\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$		

where

$$\begin{aligned}
&I_\pm = T_1 \pm \mathrm{i}T_2,\ I_3 = T_3 \\
&U_\pm = T_6 \pm \mathrm{i}T_7,\ U_3 = \frac{\sqrt{3}}{2}T_8 - \frac{1}{2}T_3 = \frac{3}{4}Y - \frac{1}{2}I_3 \\
&V_\pm = T_4 \pm \mathrm{i}T_5,\ V_3 = \frac{\sqrt{3}}{2}T_8 + \frac{1}{2}T_3 = \frac{3}{4}Y + \frac{1}{2}I_3 \\
&Y = \frac{2}{\sqrt{3}}T_8 = \frac{1}{2}(B+S), \quad B \text{ is baryon number,} \\
&S \text{ is the quantum number ``strangeness''}
\end{aligned} \tag{5.1.20}$$

$$C_3 = 2\left(d_{ijk} + \mathrm{i}f_{ijk}\right)T_iT_jT_k \tag{5.1.21}$$

Using the values of f_{ijk} in Table 5.1, it is easy to prove that

$$\begin{aligned}
[I_+, I_-] &= 2I_3 \\
[U_+, U_-] &= 2U_3 \\
[V_+, V_-] &= 2V_3
\end{aligned} \tag{5.1.22}$$

Using (I_+, I_-, I_3) as generators, we can construct an SU(2) subgroup, called I group. The subgroup orthogonal to I group is the U(1) group constructed by the generator Y. Another SU(2) subgroup constructed by the generators (U_+, U_-, U_3) is called U group, the subgroup orthogonal to it is the U(1) subgroup Q constructed by the generator $Q = I_3 + \frac{Y}{2} = T_3 + \frac{1}{\sqrt{3}}T_8$. The SU(2) subgroup generated by the generators (V_+, V_-, V_3) is called V group, the U(1) subgroup orthogonal to it is called X group which is generated by the generator

$$X = \frac{1}{\sqrt{3}}T_8 - T_3 = \frac{1}{3}\begin{pmatrix} -1 & & \\ & 2 & \\ & & -1 \end{pmatrix}.$$

We will come back to these when discussing the quark model.

5.2. SU(3) Quark Model, the SU(3) Flavor Wave Functions for Mesons and Baryons

In 1964, Gell-Mann and Zweig proposed SU(3) quark model [5] based on the SU(3) Eightfold Way. They assumed that the more elementary building blocks of hadrons are “quarks”. Quarks have three different “flavor” that are denoted by u (up), d (down), and s (strange), respectively. The u, d, s quarks fill the fundamental representation $\underline{3}$ of the SU(3) group. Mesons consist of a pair of quark and antiquark, while baryons consist of three quarks. We denote the u, d, s quarks by q, while the antiquarks $\bar{\mathrm{u}}$, $\bar{\mathrm{d}}$, $\bar{\mathrm{s}}$ by $\bar{\mathrm{q}}$. Then mesons can be expressed as $\mathrm{q}\bar{\mathrm{q}}$, baryons as qqq. $\bar{\mathrm{q}}(\bar{\mathrm{u}}, \bar{\mathrm{d}}, \bar{\mathrm{s}})$ fills the SU(3) complex conjugate representation $\underline{3^*}$. Note that Gell-Mann called them quarks, but Zweig called them aces. Later, we only used the name “quarks”.

In order to solve the difficulty in constructing baryons for the Sakata model, a bold assumption is proposed in the quark model, that is, the baryon numbers of the u, d, s quarks are taken to be fractional number 1/3, and their electric charges are taken to be (in unit of the proton electric charge) fractional numbers 2/3, −1/3, −1/3, respectively. This bold assumption is critical for the success of the quark model. In order to construct the strange baryons, they assumed the strangeness quantum numbers $s = 0$ for the quarks u and d, $s = -1$ for the quark s. In Table 5.2, we have listed all the quark quantum numbers.

Table 5.2. u, d, s quantum numbers.

	I	I_3	B	S	Y	Q
u	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{3}$	0	$\frac{1}{3}$	$\frac{2}{3}$
d	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{3}$	0	$\frac{1}{3}$	$-\frac{1}{3}$
s	0	0	$\frac{1}{3}$	−1	$-\frac{2}{3}$	$-\frac{1}{3}$

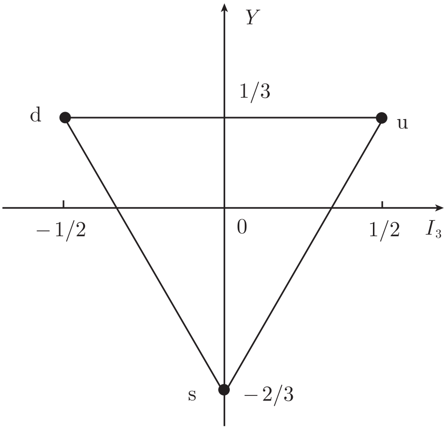


Fig. 5.4. The weight diagram for u, d, s quarks.

Assuming hypercharge Y as the vertical axis, and I_3 as the horizontal axis, we can draw the weight diagram for u, d, s quarks in Fig. 5.4.

We can express u, d, s quarks as the states of $|I, I_3\rangle$ and the column vectors

$$|\mathrm{u}\rangle = \left|\frac{1}{2},\frac{1}{2}\right\rangle = \begin{pmatrix}1\\0\\0\end{pmatrix}, \quad |\mathrm{d}\rangle = \left|\frac{1}{2},-\frac{1}{2}\right\rangle = \begin{pmatrix}0\\1\\0\end{pmatrix}, \quad |\mathrm{s}\rangle = |0,0\rangle = \begin{pmatrix}0\\0\\1\end{pmatrix} \tag{5.2.1}$$

Using $I_3 = T_3, Y = \frac{2}{\sqrt{3}}T_8$, $Q = T_3 + \frac{1}{\sqrt{3}}T_8$ to act on Eq. (5.2.1), we can obtain all the quantum numbers in Table 5.2. From Eqs. (5.1.20) and (5.2.1), we get

$$I_+ |\mathrm{d}\rangle = |\mathrm{u}\rangle, \quad I_- |\mathrm{u}\rangle = |\mathrm{d}\rangle \tag{5.2.2}$$

Thus, $I_\pm$ can make d ↔ u mutual transformation: the $U_\pm$ in U spin can make d ↔ s; and the $V_\pm$ in V spin can make s ↔ u. The results for I spin, U spin, and V spin acting on u, d, s quarks are illustrated in Fig. 5.5.

Note that $q^i = (\mathrm{u}, \mathrm{d}, \mathrm{s})$ transforms according to u under the SU(3) transformation, i.e. $q^i \to q'^i \to u^i_j q^j$, so q^i is a 3-dimensional vector 和. While $\bar{q}_i = (\bar{\mathrm{u}}, \bar{\mathrm{d}}, \bar{\mathrm{s}})$ transforms according to u^+, i.e. $\bar{q}_i \to \bar{q}'_i \to (u^+)^j_i\, \bar{q}_j$, so $\bar{q}_i$ is also a 3-dimensional vector. The only difference is that q^i belongs to representation $\underline{3}$ and $\bar{q}_i$ belongs to $\underline{3^*}$. The representations $\underline{3}$ and $\underline{3^*}$ are isomorphic, but not equivalent, so there is no similarity transformation S which makes $u^* = SuS^{-1}$.

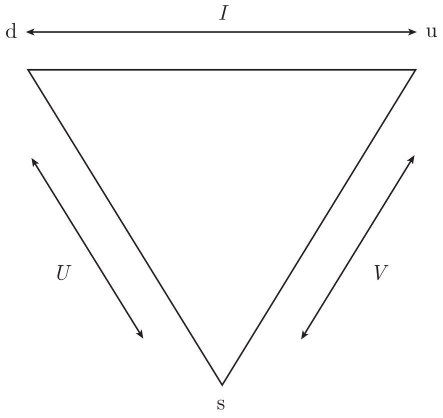


Fig. 5.5. I, U, V spins.

The wave functions for mesons and baryons include four parts:

$$\Psi = \psi_{\text{space}} \cdot \psi_{\text{spin}} \cdot \psi_{\text{flavor}} \cdot \psi_{\text{color}} \tag{5.2.3}$$

where the space part ψ_{space} is related to the space coordinates x only, the spin part ψ_{spin} is only connected to the spin, ψ_{flavor} is only relative with favor quantum numbers, while ψ_{color} will only do something with the color degrees of freedom. In what follows, we first discuss ψ_{flavor}.

5.2.1. *The flavor wave functions for the pseudoscalar meson octet and singlet*

According to quark model, mesons consist of $q\bar{q}$. In the scheme of SU(3) symmetry, $q\bar{q} = 3 \times 3^*$.

In the language of Young tableau, the representations $\underline{3}$ and $\underline{3^*}$ are expressed as $\square$, and $\begin{array}{|c|}\hline \ \\ \hline \ \\ \hline \end{array}$, respectively. According to the decomposition procedure described in Sec. 5.1.1,

$$\square \otimes \begin{array}{|c|}\hline \ \\ \hline \ \\ \hline \end{array} = \square \otimes \begin{array}{|c|}\hline a \\ \hline b \\ \hline \end{array} = \begin{array}{|c|c|}\hline \ & a \\ \hline b \\ \cline{1-1} \end{array} + \begin{array}{|c|}\hline \ \\ \hline a \\ \hline b \\ \hline \end{array} = \underline{8} + \underline{1} \tag{5.2.4}$$

The dimensions of the last two tableaux are easy to calculate using Eq. (5.1.6). So mesons fill octet $\underline{8}$ and singlet $\underline{1}$. The detail flavor wave function ψ_{flavor} can be obtained from tensor analysis.

Put

$$M'^i_j = q^i \bar{q}_j \tag{5.2.5}$$

obviously $M'^i_i = q^i \bar{q}_i \neq 0$ (the repeated index means summation). According to Eq. (5.1.11), M'^i_j is the reducible tensor. Defining a new tensor

$$M^i_j = q^i \bar{q}_j - \frac{1}{3}\delta^i_j q^k \bar{q}_k \equiv M'^i_j - \frac{1}{3}M'^k_k \delta^i_j \tag{5.2.6}$$

we can easily see that the new tensor fulfils the traceless condition

$$M^i_i = 0 \tag{5.2.7}$$

thus M^i_j is an irreducible tensor. Note that $M'^i_i = q^i \bar{q}_i \neq 0$ is an SU(3) invariant, so it belongs to a singlet $\underline{1}$.

Now M^i_j is traceless tensor, so it has only eight independent components (the traceless condition eliminates one independent component, so the number of independent components for M^i_j are $9 - 1 = 8$). Therefore the traceless tensor M^i_j forms $\underline{8}$-dimensional representation of the SU(3) group.

Using $q^i = (\mathrm{u}, \mathrm{d}, \mathrm{s})$ and $\bar{q}_j = (\bar{\mathrm{u}}, \bar{\mathrm{d}}, \bar{\mathrm{s}})$, it is easy to obtain

$$M^1_1 = q^1 \bar{q}_1 - \frac{1}{3} q^k \bar{q}_k \equiv \mathrm{u}\bar{\mathrm{u}} - \frac{1}{3}(\mathrm{u}\bar{\mathrm{u}} + \mathrm{d}\bar{\mathrm{d}} + \mathrm{s}\bar{\mathrm{s}}) = \frac{1}{3}(2\mathrm{u}\bar{\mathrm{u}} - \mathrm{d}\bar{\mathrm{d}} - \mathrm{s}\bar{\mathrm{s}}) \tag{5.2.8}$$

$$M^1_2 = \mathrm{u}\bar{\mathrm{d}}, \quad M^1_3 = \mathrm{u}\bar{\mathrm{s}}$$

etc., so we obtain

$$M^i_j = \begin{pmatrix} \dfrac{2\mathrm{u}\bar{\mathrm{u}} - \mathrm{d}\bar{\mathrm{d}} - \mathrm{s}\bar{\mathrm{s}}}{3} & \mathrm{u}\bar{\mathrm{d}} & \mathrm{u}\bar{\mathrm{s}} \\ \mathrm{d}\bar{\mathrm{u}} & \dfrac{2\mathrm{d}\bar{\mathrm{d}} - \mathrm{u}\bar{\mathrm{u}} - \mathrm{s}\bar{\mathrm{s}}}{3} & \mathrm{d}\bar{\mathrm{s}} \\ \mathrm{s}\bar{\mathrm{u}} & \mathrm{s}\bar{\mathrm{d}} & \dfrac{2\mathrm{s}\bar{\mathrm{s}} - \mathrm{u}\bar{\mathrm{u}} - \mathrm{d}\bar{\mathrm{d}}}{3} \end{pmatrix} \tag{5.2.9}$$

The singlet is given by

$$q^i \bar{q}_i = \frac{\mathrm{u}\bar{\mathrm{u}} + \mathrm{d}\bar{\mathrm{d}} + \mathrm{s}\bar{\mathrm{s}}}{\sqrt{3}} \equiv \eta_1 \tag{5.2.10}$$

where the $\sqrt{3}$ is the normalization factor. In the diagonal line of Eq. (5.2.9), there are only two independent states; the traceless condition eliminates one state. We now first look at the pseudoscalar mesons ($J^{\mathrm{P}} = 0^-$). According to the quark quantum numbers in Table 5.2, we have the following conditions:

$$\left.\begin{array}{c}
\mathrm{K}^0 = \mathrm{d}\bar{\mathrm{s}}, \qquad \mathrm{K}^+ = \mathrm{u}\bar{\mathrm{s}} \\[1ex]
\pi^+ = \mathrm{u}\bar{\mathrm{d}}, \qquad \pi^0 = \dfrac{\mathrm{d}\bar{\mathrm{d}} - \mathrm{u}\bar{\mathrm{u}}}{\sqrt{2}}, \qquad \pi^- = -\mathrm{d}\bar{\mathrm{u}} \\[2ex]
\eta_8 = \dfrac{\mathrm{u}\bar{\mathrm{u}} + \mathrm{d}\bar{\mathrm{d}} - 2\mathrm{s}\bar{\mathrm{s}}}{\sqrt{6}}, \\[2ex]
\mathrm{K}^- = -\mathrm{s}\bar{\mathrm{u}}, \qquad \mathrm{K}^0 = \mathrm{s}\bar{\mathrm{d}}
\end{array}\right\} \tag{5.2.11}$$

where $(-\bar{\mathrm{u}})$ is $I_3 = -1/2$ state. From the isospin symmetry in Chapter 4, we know that only $\begin{pmatrix} \mathrm{u} \\ \mathrm{d} \end{pmatrix}$ and $\begin{pmatrix} \bar{\mathrm{d}} \\ -\bar{\mathrm{u}} \end{pmatrix}$ can have common isospin transformation property, i.e. only $(-\bar{\mathrm{u}})$ is $I_3 = -1/2$ antiquark state with proper isospin transformation property. So by Eq. (5.2.11), there must be a minus sign in front of π^-, K^-, i.e. we use $(-\bar{\mathrm{u}})$ in constructing π^-, K^-. For π^0, we use $\frac{\mathrm{d}\bar{\mathrm{d}} - \mathrm{u}\bar{\mathrm{u}}}{\sqrt{2}}$.

In fact, the two independent quantities in the diagonal line of Eq. (5.2.9) can be chosen as follows:

$$\begin{aligned}
M_2^2 - M_1^1 &= \mathrm{d}\bar{\mathrm{d}} - \mathrm{u}\bar{\mathrm{u}} \sim \pi^0 \\
M_2^2 + M_1^1 &= \frac{\mathrm{u}\bar{\mathrm{u}} + \mathrm{d}\bar{\mathrm{d}} - 2\mathrm{s}\bar{\mathrm{s}}}{\sqrt{6}} \sim \eta_8
\end{aligned} \tag{5.2.12}$$

Here, for the first quantity $I = 1$, $I_3 = 0$ — it can be chosen as π^0 and the normalization factor is $\sqrt{2}$, while, the second quantity $I = 0$, $I_3 = 0$ — it can be chosen as η_8 and the normalization factor is $\sqrt{6}$. So Eq. (5.2.9) can be rewritten as

$$M_j^i = \begin{pmatrix}
\dfrac{\eta_8}{\sqrt{6}} - \dfrac{\pi^0}{\sqrt{2}} & \pi^+ & \mathrm{K}^+ \\[2ex]
-\pi^- & \dfrac{\eta_8}{\sqrt{6}} + \dfrac{\pi^0}{\sqrt{2}} & \mathrm{K}^0 \\[2ex]
-\mathrm{K}^- & \bar{\mathrm{K}}^0 & -\dfrac{2\eta_8}{\sqrt{6}}
\end{pmatrix} \tag{5.2.13}$$

Further η_1 in Eq. (5.2.10) and η_8 in Eq. (5.2.11) are all particles of $I = 0$, $Y = 0$. If η_1 is SU(3) singlet, then η_8 belongs to SU(3) octet. If SU(3) symmetry is rigorous, then η_1 and η_8 are not relevant to each other. If in fact SU(3) symmetry is weakly broken, this induces the mixing of η_1 and

η_8. The observed η, η' mesons in experiments are the products after the mixing of η_1 and η_8. Put the mixing angle as θ, then

$$\begin{pmatrix} \eta \\ \eta' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \eta_8 \\ \eta_1 \end{pmatrix} \tag{5.2.14}$$

Thus we get

$$\begin{aligned} \eta &= \eta_8 \cos\theta - \eta_1 \sin\theta \\ \eta' &= \eta_8 \sin\theta + \eta_1 \cos\theta \end{aligned} \tag{5.2.15}$$

The mixing angle can be calculated from the measured meson masses; this needs to use mass formula. The calculation method will be given when discussing mass formula. At present, the calculated θ has two different values: $\theta = -11.5°$ for mass squired formula, while $\theta = -24.6°$ for linear mass formula [7].

5.2.2. *Flavor wave functions for the vector meson octet and singlet*

Making a replacement in the flavor wave functions for the pseudoscalar mesons in Sec. 5.2.1: $\pi \to \rho$, $\mathrm{K} \to \mathrm{K}^*$, $\eta_8 \to \omega_8$, $\eta_1 \to \omega_1$ we get immediately the flavor wave functions for the vector meson octet and singlet. Although the flavor wave functions for the pseudoscalar mesons and vector masons have the same quark structure, the spin structure is different. For pseudoscalar meson, the spins of q and $\bar{\mathrm{q}}$ are antiparallel, so the total spin is 0, while in vector meson the spins of q and $\bar{\mathrm{q}}$ are parallel, so the total spin is 1. Certainly the mixing of $\eta_8 - \eta_1$ and that of $\omega_8 - \omega_1$ are also different.

Similarly, the mixing of $\omega_1 - \omega_8$ can be expressed as

$$\begin{aligned} \phi &= \omega_8 \cos\theta - \omega_1 \sin\theta \\ \omega &= \omega_8 \sin\theta + \omega_1 \cos\theta \end{aligned} \tag{5.2.16}$$

For "ideal mixing", take $\cos\theta = \sqrt{\frac{2}{3}}$, $\sin\theta = \sqrt{\frac{1}{3}}$, that is, $\theta = 35°$, then

$$\phi = \frac{\mathrm{u\bar{u}} + \mathrm{d\bar{d}} - 2\mathrm{s\bar{s}}}{\sqrt{6}} \cdot \sqrt{\frac{2}{3}} - \frac{\mathrm{u\bar{u}} + \mathrm{d\bar{d}} + \mathrm{s\bar{s}}}{\sqrt{3}} \sqrt{\frac{1}{3}} = -\mathrm{s\bar{s}}$$

$$\omega = \frac{\mathrm{u\bar{u}} + \mathrm{d\bar{d}} - 2\mathrm{s\bar{s}}}{\sqrt{6}} \cdot \frac{1}{\sqrt{3}} + \frac{\mathrm{u\bar{u}} + \mathrm{d\bar{d}} + \mathrm{s\bar{s}}}{\sqrt{3}} \sqrt{\frac{2}{3}} = \frac{\mathrm{u\bar{u}} + \mathrm{d\bar{d}}}{\sqrt{2}}$$

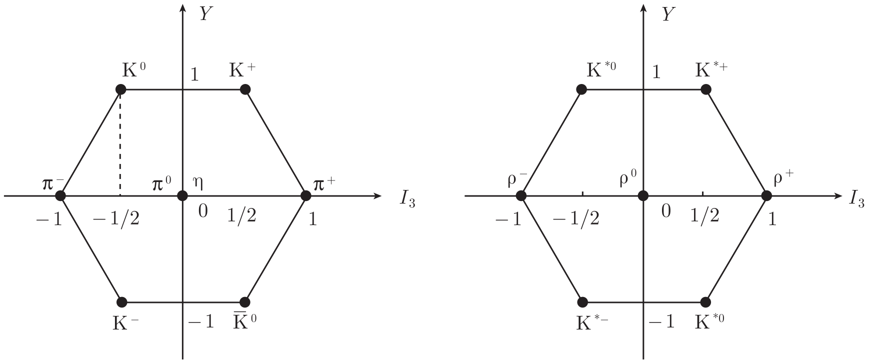


Fig. 5.6. Weight diagram for the pseudoscalar and vector meson octets.

i.e.

$$\begin{aligned}\phi &= -\mathrm{s\bar{s}} \\ \omega &= \frac{1}{\sqrt{2}}(\mathrm{u\bar{u}} + \mathrm{d\bar{d}})\end{aligned} \tag{5.2.17}$$

In the literature usually people take

$$\begin{aligned}\phi &= \mathrm{s\bar{s}} \\ \omega &= \frac{1}{\sqrt{2}}(\mathrm{u\bar{u}} + \mathrm{d\bar{d}})\end{aligned} \tag{5.2.18}$$

ϕ in Eq. (5.2.18) and ϕ in Eq. (5.2.17) have phase difference 180°. This does not affect physics.

The $\omega - \phi$ mixing angle measured in experiments is very close to the ideal mixing one. For the mass squared fit formula $\theta = 38.7°$, while for the linear mass formula $\theta = 36.0°$ [7]. The weight diagrams for the pseudoscalar and vector meson octets are shown in Fig. 5.6.

5.2.3. *Flavor wave functions for baryon octet and decuplet*

Baryons are consist of three quarks. From Sec. 5.1.1 we know that

$$\begin{aligned} \mathrm{q}^a\mathrm{q}^b\mathrm{q}^c = 3\otimes 3\otimes 3 = \boxed{a}\otimes\boxed{b}\otimes\boxed{c} &= \left(\begin{array}{|c|c|}\hline a & b \\ \hline \end{array} + \begin{array}{|c|}\hline a \\ \hline b \\ \hline \end{array}\right)\otimes\boxed{c} \\ &\qquad\quad \underline{6} + \underline{3^*} \\ &= \begin{array}{|c|c|c|}\hline a & b & c \\ \hline \end{array} + \begin{array}{|c|c|}\hline a & b \\ \hline c \\ \cline{1-1} \end{array} + \begin{array}{|c|c|}\hline a & c \\ \hline b \\ \cline{1-1} \end{array} + \begin{array}{|c|}\hline a \\ \hline b \\ \hline c \\ \hline \end{array} \\ &\quad\ \ \underline{10} \quad + \quad \underline{8} \quad + \quad \underline{8} \quad + \quad \underline{1} \end{aligned} \tag{5.2.19}$$

The above decomposition can be written as

$$3 \times 3 \times 3 = 10 \oplus 8 \oplus 8 \oplus 1$$

where there are two octets, but they have different symmetry property.

From the above analysis of Young tableaux, we can see that we first decompose two quark states $\mathrm{q}^a\mathrm{q}^b$ into $\underline{6}+\underline{3}^*$, i.e. $3\times 3 = \underline{6}+\underline{3}^*$, then make the third quark to join in. In order to write down the quark structure of baryon wave functions, we use again tensor analysis.

Because quarks transform according to SU(3) fundamental representation $\underline{3}$, so the tensor $T^{ab} = \mathrm{q}^a\mathrm{q}^b$ can be reduced to $6+3^*$, i.e.

$$\begin{array}{cc} \begin{array}{|c|c|}\hline a & b \\ \hline\end{array} & + \begin{array}{|c|}\hline a \\ \hline b \\ \hline\end{array} \\ \\ \underline{6} & + \underline{3}^* \end{array} \tag{5.2.20}$$

while the tensor of $\underline{3^*}$, i.e. $A_i = \varepsilon_{iab}T^{ab}$, is antisymmetric with respect to the indices a, b. This can be seen from the Young tableaux above

$$\begin{aligned} A_1 &= \frac{1}{\sqrt{2}}\left(T^{23} - T^{32}\right) \\ A_2 &= \frac{1}{\sqrt{2}}\left(T^{31} - T^{13}\right) \\ A_3 &= \frac{1}{\sqrt{2}}\left(T^{12} - T^{21}\right) \end{aligned} \tag{5.2.21}$$

where $\frac{1}{\sqrt{2}}$ is the normalization factor.

From Eq. (5.2.20), the 6-dimensional tensor is symmetric with respect to a, b, so

$$\begin{aligned} \underline{6} = S^{ab} &= \frac{1}{\sqrt{2}}\left(T^{ab} + T^{ba}\right) \\ &= \left\{ S^{11} = \frac{T^{11}+T^{11}}{\sqrt{2}}, S^{12} = \frac{T^{12}+T^{21}}{\sqrt{2}}, \right. \\ S^{13} &= \frac{T^{13}+T^{31}}{\sqrt{2}}, S^{23} = \frac{T^{23}+T^{32}}{\sqrt{2}}, \\ S^{22} &= \left. \frac{T^{22}+T^{22}}{\sqrt{2}}, S^{33} = \frac{T^{33}+T^{33}}{\sqrt{2}} \right\} \end{aligned} \tag{5.2.22}$$

Expressing S^{ii} as $\frac{T^{ii}+T^{ii}}{\sqrt{2}}$ is a correct normalization; it can be understood as $S^{ii} = \frac{\mathrm{q}^i(1)\mathrm{q}^i(2)+\mathrm{q}^i(2)\mathrm{q}^i(1)}{\sqrt{2}}$, so

$$T^{ab} = \mathrm{q}^a\mathrm{q}^b \to \underline{6} + \underline{3}^* \to S^{ab} + A_i \tag{5.2.23}$$

$$T^{abc} = \mathrm{q}^a\mathrm{q}^b\mathrm{q}^c = \left(S^{ab} + A_i\right)\mathrm{q}^c \tag{5.2.24}$$

Now, we first discuss decomposition of $A_i\mathrm{q}^c$ (i.e. $3^* \times 3$):

$$\begin{array}{c} A_i\mathrm{q}^c = A_i\mathrm{q}^i + B^i_j \to 1 \oplus 8 \\ \downarrow \qquad \downarrow \\ \begin{array}{|c|}\hline a \\ \hline b \\ \hline c \\ \hline \end{array} \qquad \begin{array}{|c|c|}\hline a & c \\ \hline b & \multicolumn{1}{c}{} \\ \cline{1-1} \end{array} \end{array} \tag{5.2.25}$$

where the singlet

$$\frac{1}{\sqrt{3}}A_i\mathrm{q}^i = \frac{1}{\sqrt{6}}\left(\mathrm{dsu} - \mathrm{sdu} + \mathrm{sud} - \mathrm{usd} + \mathrm{uds} - \mathrm{dus}\right) \tag{5.2.26}$$

is totally antisymmetric with respect to a, b, c and $\frac{1}{\sqrt{6}}$ is the normalization factor.

$\frac{1}{2}^+$ baryon octet is given by

$$\begin{aligned} B^i_j &= A_j\mathrm{q}^i - \frac{1}{3}\delta^i_j A_k\mathrm{q}^k \\ &= \varepsilon_{jab}\mathrm{q}^a\mathrm{q}^b\mathrm{q}^i - \frac{1}{3}\delta^i_j\varepsilon_{abc}\mathrm{q}^a\mathrm{q}^b\mathrm{q}^c \end{aligned} \tag{5.2.27}$$

The detail form is as follows:

$$\left.\begin{aligned}
&B^1_1 = \mathrm{dsu} - \mathrm{sdu} - \frac{1}{3}A_kq^k \\
&B^1_2 = \mathrm{suu} - \mathrm{usu} && \to \Sigma^+ = \frac{1}{\sqrt{2}}\left(\mathrm{suu} - \mathrm{usu}\right) \\
&B^1_3 = \mathrm{udu} - \mathrm{duu} && \to \mathrm{P} = \frac{1}{\sqrt{2}}\left(\mathrm{udu} - \mathrm{duu}\right) \\
&B^2_1 = \mathrm{dsd} - \mathrm{sdd} && \to \Sigma^- = \frac{1}{\sqrt{2}}\left(\mathrm{dsd} - \mathrm{sdd}\right) \\
&B^2_2 = \mathrm{sud} - \mathrm{usd} - \frac{1}{3}A_kq^k \\
&B^2_3 = \mathrm{udd} - \mathrm{dud} && \to \mathrm{n} = \frac{1}{\sqrt{2}}\left(\mathrm{udd} - \mathrm{dud}\right) \\
&B^3_1 = \mathrm{dss} - \mathrm{sds} && \to \Xi^- = \frac{1}{\sqrt{2}}\left(\mathrm{dss} - \mathrm{sds}\right) \\
&B^3_2 = \mathrm{sus} - \mathrm{uss} && \to \Xi^0 = \frac{1}{\sqrt{2}}\left(\mathrm{sus} - \mathrm{uss}\right) \\
&B^3_3 = \mathrm{uds} - \mathrm{dus} - \frac{1}{3}A_kq^k
\end{aligned}\right\} \tag{5.2.28}$$

In Eq. (5.2.28), $1/\sqrt{2}$ is the normalization factor. Because $B^i_i = B^1_1 + B^2_2 + B^3_3 = 0$, $B^i_i\,(i = 1, 2, 3)$ have only two independent components.

Similar to the choice for meson octet in Eq. (5.2.12), we can define

$$\pi^0 \to \Sigma^0 = \frac{1}{2}\left(B_2^2 - B_1^1\right) = \frac{1}{2}\left(\mathrm{sud} - \mathrm{usd} - \mathrm{dsu} + \mathrm{sdu}\right) \tag{5.2.29}$$

$$\begin{aligned}\eta_8 \to \Lambda^0 &= \frac{\sqrt{3}}{2}\left(B_2^2 + B_1^1\right) = -\frac{\sqrt{3}}{2}\mathrm{B}_3^3 \\ &= \frac{1}{\sqrt{12}}\left\{\left(\mathrm{ds} - \mathrm{sd}\right)\mathrm{u} + \left(\mathrm{su} - \mathrm{us}\right)\mathrm{d} - 2\left(\mathrm{ud} - \mathrm{du}\right)\mathrm{s}\right\}\end{aligned} \tag{5.2.30}$$

The normalization factors in Eqs. (5.2.29) and (5.2.30) are different to that of $\pi^0 - \eta_8$, the reason is that the terms for the quark structure are different for the two cases. Obviously

$$\frac{\Lambda^0}{\sqrt{6}} - \frac{\Sigma^0}{\sqrt{2}} = \frac{-1}{2\sqrt{2}}\left(B_3^3 + B_2^2 - B_1^1\right) = \frac{1}{\sqrt{2}}B_1^1 \tag{5.2.31}$$

so we get

$$\frac{1}{\sqrt{2}}\left(B_j^i\right) = \begin{pmatrix} \dfrac{\Lambda^0}{\sqrt{6}} - \dfrac{\Sigma^0}{\sqrt{2}} & \Sigma^+ & \mathrm{p} \\ \Sigma^- & \dfrac{\Lambda^0}{\sqrt{6}} + \dfrac{\Sigma^0}{\sqrt{2}} & \mathrm{n} \\ \Xi^- & \Xi^0 & -\dfrac{2\Lambda^0}{\sqrt{6}} \end{pmatrix} \tag{5.2.32}$$

Comparing with Eq. (5.2.13), there are no minus sign in front of Σ^-, Ξ^-. This does not matter and the phase difference of 180° in wave function does not affect physics.

In summary the $\frac{1}{2}^+$ baryon assignment for B_j^i is

$$\left.\begin{aligned} &\Sigma^+ = \frac{1}{\sqrt{2}}\left(\mathrm{suu} - \mathrm{usu}\right) \\ &\mathrm{p} = \frac{1}{\sqrt{2}}\left(\mathrm{udu} - \mathrm{duu}\right) \\ &\Sigma^- = \frac{1}{\sqrt{2}}\left(\mathrm{dsd} - \mathrm{sdd}\right) \\ &\mathrm{n} = \frac{1}{\sqrt{2}}\left(\mathrm{udd} - \mathrm{dud}\right) \\ &\Xi^- = \frac{1}{\sqrt{2}}\left(\mathrm{dss} - \mathrm{sds}\right) \\ &\Xi^0 = \frac{1}{\sqrt{2}}\left(\mathrm{sus} - \mathrm{uss}\right) \\ &\Lambda^0 = \frac{1}{\sqrt{12}}\left(\mathrm{dsu} - \mathrm{sdu} + \mathrm{sud} - \mathrm{usd} + 2\mathrm{uds} - 2\mathrm{dus}\right) \\ &\Sigma^0 = \frac{1}{2}\left(\mathrm{sud} - \mathrm{usd} - \mathrm{dsu} + \mathrm{sdu}\right) \end{aligned}\right\} \tag{5.2.33}$$

Note that the baryon octet in Eq. (5.2.27) corresponds to the Young tableau $\begin{array}{|c|c|}\hline a & c \\ \hline b \\ \cline{1-1}\end{array}$ in Eq. (5.2.19), i.e. the symbols a and b are antisymmetric. There is another octet, $\begin{array}{|c|c|}\hline a & b \\ \hline c \\ \cline{1-1}\end{array}$, where a, b are symmetric. They are all come from the decomposition of $\underline{6} \times \underline{3}$. In fact, we also can choose

$$B'^{b}_{j} = \varepsilon_{ajc} S^{ab} \mathrm{q}^c \tag{5.2.34}$$

B'^{b}_{j} is an irreducible tensor and belongs to 8-dimensional representation. This can be seen from $B'^{j}_{j} = \varepsilon_{ajc} S^{aj} \mathrm{q}^c = 0$ (cf. Eq. (5.1.11)). In the right-hand side of Eq. (5.2.34), different position of the lower index j, i.e. ε_{ajc}, ε_{jac}, ε_{acj}, just has overall sign difference; it does not affect physics.

From Eq. (5.2.34), we obtain (using the expression of S^{ab} in Eq. (5.2.22))

$$\begin{aligned}
B'^{1}_{2} &= \varepsilon_{a2c} S^{a1} \mathrm{q}^c = S^{11}\mathrm{q}^3 - S^{31}\mathrm{q}^1 = \frac{\mathrm{uus}+\mathrm{uus}}{\sqrt{2}} - \frac{\mathrm{suu}+\mathrm{usu}}{\sqrt{2}} \\
&= \frac{1}{\sqrt{2}}\left(2\mathrm{uus} - \mathrm{suu} - \mathrm{usu}\right) \\
&\to \frac{1}{\sqrt{6}}\left(2\mathrm{uus} - \mathrm{suu} - \mathrm{usu}\right) = \Sigma^+ \\
B'^{1}_{3} &= \varepsilon_{a3c} S^{a1} q^c = S^{21} q^1 - S^{11} q^2 \\
&= \frac{(\mathrm{du}+\mathrm{ud})\,\mathrm{u}}{\sqrt{2}} - \frac{(\mathrm{uu}+\mathrm{uu})\,d}{\sqrt{2}} = \frac{1}{\sqrt{2}}\left(\mathrm{duu}+\mathrm{udu}-2\mathrm{uud}\right) \\
&\to \frac{1}{\sqrt{6}}\left(\mathrm{duu}+\mathrm{udu}-2\mathrm{uud}\right) = \mathrm{p} \\
B'^{2}_{1} &= \varepsilon_{a1c} S^{a2} \mathrm{q}^c = S^{32}\mathrm{q}^2 - S^{22}\mathrm{q}^3 = \frac{1}{\sqrt{2}}(\mathrm{sd}+\mathrm{ds})\,\mathrm{d} - \frac{1}{\sqrt{2}}(\mathrm{dd}+\mathrm{dd})\,\mathrm{s} \\
&\to \frac{1}{\sqrt{6}}\left(\mathrm{sdd}+\mathrm{dsd}-2\mathrm{dds}\right) = \Sigma^- \\
B'^{2}_{3} &= \varepsilon_{a3c} S^{a2} \mathrm{q}^c = S^{22}\mathrm{q}^1 - S^{12}\mathrm{q}^2 = \frac{1}{\sqrt{2}}(\mathrm{dd}+\mathrm{dd})\,\mathrm{u} - \frac{1}{\sqrt{2}}(\mathrm{ud}+\mathrm{du})\,\mathrm{d} \\
&\to \frac{1}{\sqrt{6}}\left(2\mathrm{ddu}-\mathrm{udd}-\mathrm{dud}\right) = \mathrm{n} \\
B'^{3}_{1} &= \varepsilon_{a1c} S^{a3} \mathrm{q}^c = S^{33}\mathrm{q}^2 - S^{23} q^3 \\
&\to \frac{1}{\sqrt{6}}\left(2\mathrm{ssd}-\mathrm{dss}-\mathrm{sds}\right) = \Xi^-
\end{aligned}$$

$$\begin{aligned}
B'^{3}_{2} &= \varepsilon_{a2c} S^{a3} \mathrm{q}^{c} = S^{13}\mathrm{q}^{3} - S^{33}\mathrm{q}^{1} \\
&= \frac{1}{\sqrt{2}}\,(\mathrm{us}+\mathrm{su})\,\mathrm{s} - \frac{1}{\sqrt{2}}\,(\mathrm{ss}+\mathrm{ss})\,\mathrm{u} \\
&\to \frac{1}{\sqrt{6}}\,(\mathrm{uss}+\mathrm{sus}-2\mathrm{ssu}) = \Xi^0
\end{aligned}$$

$$B'^{1}_{1} = \varepsilon_{a1c} S^{a1} \mathrm{q}^{c} = S^{31}\mathrm{q}^{2} - S^{21}\mathrm{q}^{3} = \frac{1}{\sqrt{2}}\,(\mathrm{sud}+\mathrm{usd}-\mathrm{dus}-\mathrm{uds})$$

$$B'^{2}_{2} = \varepsilon_{a2c} S^{a2} \mathrm{q}^{c} = S^{12}\mathrm{q}^{3} - S^{32}\mathrm{q}^{1} = \frac{1}{\sqrt{2}}\,(\mathrm{uds}+\mathrm{dus}-\mathrm{sdu}-\mathrm{dsu})$$

$$B'^{3}_{3} = \varepsilon_{a3c} S^{a3} \mathrm{q}^{c} = S^{23}\mathrm{q}^{1} - S^{13}\mathrm{q}^{2} = \frac{1}{\sqrt{2}}\,(\mathrm{dsu}+\mathrm{sdu}-\mathrm{usd}-\mathrm{sud})$$

Obviously $B'^{1}_{1} + B'^{2}_{2} + B'^{3}_{3} = 0$.

Comparing with Eq. (5.2.12) we can define

$$\begin{aligned}
\Sigma^0 &= \frac{1}{\sqrt{2}}(B'^{2}_{2} - B'^{1}_{1}) \to \frac{1}{\sqrt{12}}\,(2\mathrm{uds}+2\mathrm{dus}-\mathrm{sdu}-\mathrm{dsu}-\mathrm{sud}-\mathrm{usd}) \\
\Lambda^0 &= \frac{\sqrt{3}}{2}(B'^{2}_{2} + B'^{1}_{1}) = -\frac{\sqrt{3}}{2}B'^{3}_{3} \\
&\Rightarrow \Lambda^0 = \frac{1}{2}\,(\mathrm{dsu}-\mathrm{usd}+\mathrm{sdu}-\mathrm{sud})
\end{aligned}$$

so the other baryon octet B'^{i}_{j} can be expressed as

$$\begin{aligned}
\Sigma^+ &= \frac{1}{\sqrt{6}}\,(2\mathrm{uus}-\mathrm{suu}-\mathrm{usu}) \\
\mathrm{p} &= \frac{1}{\sqrt{6}}\,(\mathrm{duu}+\mathrm{udu}-2\mathrm{uud}) \\
\Sigma^- &= \frac{1}{\sqrt{6}}\,(\mathrm{sdd}+\mathrm{dsd}-2\mathrm{dds}) \\
n &= \frac{1}{\sqrt{6}}\,(2\mathrm{ddu}-\mathrm{udd}-\mathrm{dud}) \\
\Xi^- &= \frac{1}{\sqrt{6}}\,(2\mathrm{ssd}-\mathrm{dss}-\mathrm{sds}) \\
\Xi^0 &= \frac{1}{\sqrt{6}}\,(\mathrm{uss}+\mathrm{sus}-2\mathrm{ssu}) \\
\Lambda^0 &= \frac{1}{2}\,(\mathrm{dsu}-\mathrm{usd}+\mathrm{sdu}-\mathrm{sud}) \\
\Sigma^0 &= \frac{1}{\sqrt{12}}\,(2\mathrm{uds}+2\mathrm{dus}-\mathrm{sdu}-\mathrm{dsu}-\mathrm{sud}-\mathrm{usd})
\end{aligned} \tag{5.2.35}$$

B^i_j and ${B'}^i_j$ are equivalent for describing $\left(\frac{1}{2}\right)^+$ baryon octet, i.e. Eqs. (5.2.33) and (5.2.35) are equivalent description. These wave function structures are the same as given in [8] using spin coupling except for some global phase difference (minus sign); this does not affect physics.

From Eq. (5.2.19), we see that there is another totally symmetric tensor $\underline{10}$ decuplet left which is not filled up. We just can fill it up by $\left(\frac{3}{2}\right)^+$ baryon decuplet.

The totally symmetrized tensor $T^{abc} = \mathrm{q}^a\mathrm{q}^b\mathrm{q}^c$ can be written as $D^{abc} = T^{\{abc\}}$. $\{abc\}$ denotes the totally symmetrization with respect to a, b, c. Then we have

$$\begin{aligned}
\Omega^- &= D^{333} = T^{333} = \mathrm{sss} \\
\Xi^{*0} &= D^{133} = \frac{1}{\sqrt{3}}\left(T^{133} + T^{313} + T^{331}\right) = \frac{1}{\sqrt{3}}(\mathrm{uss} + \mathrm{sus} + \mathrm{ssu}) \\
\Xi^{*-} &= D^{233} = \frac{1}{\sqrt{3}}\left(T^{233} + T^{323} + T^{332}\right) = \frac{1}{\sqrt{3}}(\mathrm{dss} + \mathrm{sds} + \mathrm{ssd}) \\
\Sigma^{*+} &= D^{311} = \frac{1}{\sqrt{3}}\left(T^{311} + T^{131} + T^{113}\right) = \frac{1}{\sqrt{3}}(\mathrm{suu} + \mathrm{usu} + \mathrm{uus}) \\
\Sigma^{*0} &= D^{312} = \frac{1}{\sqrt{6}}\left(T^{312} + T^{231} + T^{123} + T^{321} + T^{213} + T^{132}\right) \\
&= \frac{1}{\sqrt{6}}(\mathrm{sud} + \mathrm{dsu} + \mathrm{uds} + \mathrm{sdu} + \mathrm{dus} + \mathrm{usd}) \\
\Sigma^{*-} &= D^{322} = \frac{1}{\sqrt{3}}\left(T^{322} + T^{232} + T^{223}\right) = \frac{1}{\sqrt{3}}(\mathrm{sdd} + \mathrm{dsd} + \mathrm{dds}) \\
\Delta^{++} &= D^{111} = T^{111} = \mathrm{uuu} \\
\Delta^{+} &= D^{112} = \frac{1}{\sqrt{3}}\left(T^{112} + T^{121} + T^{211}\right) = \frac{1}{\sqrt{3}}(\mathrm{uud} + \mathrm{udu} + \mathrm{duu}) \\
\Delta^{0} &= D^{122} = \frac{1}{\sqrt{3}}\left(T^{122} + T^{221} + T^{212}\right) = \frac{1}{\sqrt{3}}(\mathrm{udd} + \mathrm{dud} + \mathrm{ddu}) \\
\Delta^{-} &= D^{222} = T^{222} = \mathrm{ddd}
\end{aligned} \tag{5.2.36}$$

Putting $\left(\frac{1}{2}\right)^+$ baryon octet and $\left(\frac{3}{2}\right)^+$ baryon decuplet on $I_3 - Y$ weight diagram, we obtain Fig. 5.7.

Note, the different order of quark flavors in Ξ^{*0} baryon etc. can be characterized by different quark momentum (or other quantum numbers), for example permutation of $\vec{p}, \vec{p}\,', \vec{p}\,''$. When exchanging flavor quantum numbers (flavor exchange), the order of $\vec{p}, \vec{p}\,', \vec{p}\,''$ keeps unchanged.

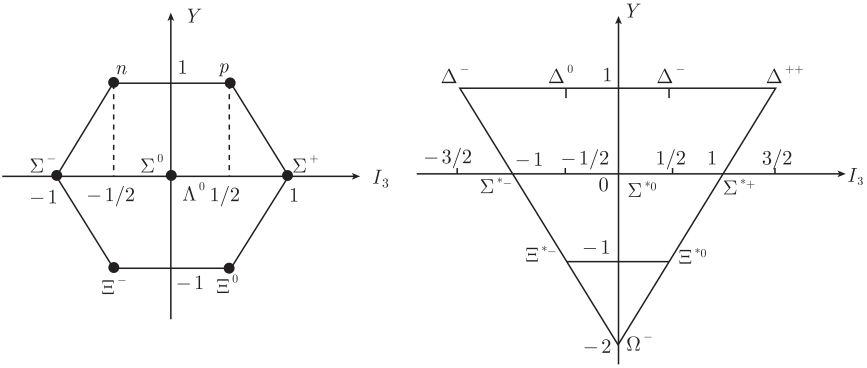


Fig. 5.7. $\left(\frac{1}{2}\right)^+$ and $\left(\frac{3}{2}\right)^+$ baryon weight diagrams.

For example

$$\begin{aligned}\mathrm{uss}+\mathrm{sus}+\mathrm{ssu} &= u\left(\vec{p}\right)s\left(\vec{p}\,'\right)s\left(\vec{p}\,''\right)+s\left(\vec{p}\right)u\left(\vec{p}\,'\right)s\left(\vec{p}\,''\right)\\ &\quad +s\left(\vec{p}\right)s\left(\vec{p}\,'\right)u\left(\vec{p}\,''\right)\\ &= D^{133}\end{aligned} \tag{5.2.37}$$

When exchanging 1, 3 (flavor exchange), the first two indices of D^{133}, i.e. 1, 3, interchange, but we do not exchange other quantum numbers (like momentum, spin, etc.), so the order of $\vec{p}$, $\vec{p}\,'$, $\vec{p}\,''$ keeps unchanged

$$D^{313}=s\left(\vec{p}\right)u\left(\vec{p}\,'\right)s\left(\vec{p}\,''\right)+u\left(\vec{p}\right)s\left(\vec{p}\,'\right)s\left(\vec{p}\,''\right)+s\left(\vec{p}\right)s\left(\vec{p}\,'\right)u\left(\vec{p}\,''\right)=D^{313} \tag{5.2.38}$$

That is, the third term in Eq. (5.2.37) does not change, the first two terms 1, 3 interchange, and the result is not changed. Certainly

$$\vec{p}+\vec{p}\,'+\vec{p}\,''=0 \tag{5.2.39}$$

5.3. Color Degree of Freedom

We have both theoretical and experimental reasons which demand quarks to have a new degree of freedom color. Quarks need to have three different color, for instance, red, green, and blue. Of course color is only a kind of quantum number, no matter what name it is. In the following, we list all the reasons which need color.

5.3.1. *The relationship of baryon spin and statistics*

According to Eq. (5.2.3), baryon wave function consists of four different parts

$$\Psi = \psi_{\text{space}} \cdot \psi_{\text{spin}} \cdot \psi_{\text{flavor}} \cdot \psi_{\text{color}} \tag{5.3.1}$$

A reasonable assumption is that quarks in hadrons are in orbital s wave. For example, $(\frac{3}{2})^+$ baryon Ω^- consists of three s quarks sss, every s quark is in orbital s wave of the common potential. So the space wave function ψ_{space} is symmetric. Because the total spin is $\frac{3}{2}$, the three quark spins must take towards the same direction, so the total spin wave function ψ_{spin} is also symmetric. We know that ψ_{flavor} is totally symmetric for $(\frac{3}{2})^+$ baryon. Thus in Eq. (5.3.1) the ψ_{color} must be antisymmetric in order to make the total wave function of $(\frac{3}{2})^+$ baryon to obey Fermi statistics. So the property of fermion statistics demands introducing color quantum number.

Using $\alpha, \beta, \gamma, \ldots = 1, 2, 3$ to denote color indices, every quark q can be expressed as $\mathrm{q}^\alpha (\alpha = 1, 2, 3)$. Baryons are color singlet, so

$$\mathrm{q}^\alpha \mathrm{q}^\beta \mathrm{q}^\gamma \varepsilon_{\alpha\beta\gamma} \tag{5.3.2}$$

Equation (5.3.2) satisfies automatically the condition of color antisymmetry. Explicitly we can write

$$\Omega^- = \frac{1}{\sqrt{6}} \mathrm{s}^\alpha \mathrm{s}^\beta \mathrm{s}^\gamma \varepsilon_{\alpha\beta\gamma} \tag{5.3.3}$$

In general, baryons consist of three quarks $q_1 q_2 q_3$, so we write $\frac{1}{\sqrt{6}} q_1^\alpha q_2^\beta q_3^\gamma \varepsilon_{\alpha\beta\gamma}$, while mesons can be expressed as $\frac{1}{\sqrt{3}} q_1^\alpha \bar{q}_2^\alpha$, where repeated index α indicates summation.

For $\left(\frac{1}{2}\right)^+$ baryons, ψ_{space} is certainly symmetric, ψ_{color} is also totally antisymmetric color singlet. ψ_{spin} and ψ_{flavor} all have mixed symmetry, but their product $\psi_{\text{spin}} \cdot \psi_{\text{flavor}}$ is symmetric; we do not discuss it further here.

5.3.2. $\pi^0 \to \gamma\gamma$

The lowest-order Feynman diagram for π^0 decaying into 2γ is shown in Fig. 5.8. The theoretical calculation for his diagram gives the decay width as

$$\Gamma\left(\pi^0 = \gamma\gamma\right) = N_{\mathrm{c}}^2 \left(Q_{\mathrm{u}}^2 - Q_{\mathrm{d}}^2\right)^2 \frac{\alpha^2 m_{\pi^0}^3}{64\pi^3 f_{\pi^2}} \tag{5.3.4}$$

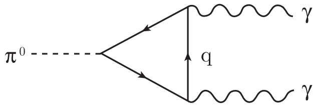


Fig. 5.8. Feynman diagram for $\pi^0 \to \gamma\gamma$.

where N_c is the number of color degrees of freedom, Q_u and Q_d are the electric charges of u and d quarks, respectively, $Q_\mathrm{u} = \frac{2}{3}, Q_\mathrm{d} = -\frac{1}{3}$, m_{π^0} is the π^0 mass, $\alpha \sim \frac{1}{137}$ is the fine structure constant, $f_\pi = 91$ MeV is the π meson decay constant. In Eq. (5.3.4) for $N_\mathrm{c} = 3$, $\Gamma\left(\pi^0 = \gamma\gamma\right) = 7.6$ eV, the result for the experimental measurements is $\mathit{\Gamma} = 7.48 \pm 0.33$ eV[7]. It matches well with the theory result for $N_\mathrm{c} = 3$. This certainly supports the existence of quark color degrees of freedom.

5.3.3. *Measurement of R value in* e$^+$e$^-$ *annihilation process*

In experiments of e$^+$e$^-$ annihilation, the definition of R is

$$R = \frac{\sigma\left(\mathrm{e^+e^-} \to \mathrm{hadrons}\right)}{\sigma\left(\mathrm{e^+e^-} \to \mu^+\mu^-\right)} \tag{5.3.5}$$

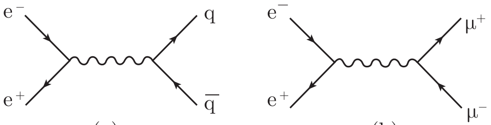
where $\sigma\left(\mathrm{e^+e^-} \to \mathrm{hadrons}\right)$ is the total cross section for e$^+$e$^-$ annihilating to any hadrons and $\sigma\left(\mathrm{e^+e^-} \to \mu^+\mu^-\right)$ is the total cross section for $\mathrm{e^+e^-} \to \mu^+\mu^-$.

e$^+$, e$^-$ have only electromagnetic interactions. Theoretically for the process of $\mathrm{e^+e^-} \to$ hadrons, e$^+$e$^-$ firstly annihilate into $\mathrm{q\bar{q}}$, $\mathrm{q\bar{q}}$ then hadronize into final state hadrons. Because until now we have not found single quark in experiments, $\mathrm{q\,\bar{q}}$ must hadronize into hadronic final states, so we can take the probability of $\mathrm{q\bar{q}} \to$ hadrons as hundred percent. Thus $\mathrm{e^+e^-} \to$ hadrons can be described by Fig. 5.9(a), while $\mathrm{e^+e^-} \to \mu^+\mu^-$ by Fig. 5.9(b).

Theoretical calculation gives

$$R = N_\mathrm{c} \sum_{i=1}^{n_\mathrm{f}} Q_i^2 \tag{5.3.6}$$

where N_c is the number of color degrees of freedom, Q_i's are the quark charges (in the unit of the proton charge $|e|$), and n_f is the number of flavors. Here n_f means the number of quark pair flavors produced in experiments.



(a) (b)

Fig. 5.9. The lowest-order Feynman diagrams for $e^+e^- \to \mu^{+-}$, $q\bar{q}$.

From experiments, we know that $m_u \sim 1.7 - 3.3$ MeV, $m_d \sim 4.1 - 5.8$ MeV, $m_s \sim 101$ MeV, $m_c \sim 1.27$ GeV, $m_b \sim 4.2 - 4.7$ GeV, $m_t \sim 172$ GeV[7]. For present e^+e^- colliders their center-of-mass energies are all less than 210 GeV, $e^+e^- \to t\bar{t}$ can not happen in experiments. So when the center-of-mass energy $\sqrt{s} < 3$ GeV, it can only produces $u\bar{u}$, $d\bar{d}$, $s\bar{s}$. Therefore, the value of R is

$$R = 3\left[Q_u^2 + Q_d^2 + Q_s^2\right] = 3\left[\left(\frac{2}{3}\right)^2 + \left(-\frac{1}{3}\right)^2 + \left(-\frac{1}{3}\right)^2\right] = 2 \qquad (5.3.7)$$

When 3 GeV $\leq \sqrt{s} \leq$ 10 GeV, u, d, s, c all can be produced, $Q_c = \frac{2}{3}$, so

$$R = 3\left[Q_u^2 + Q_d^2 + Q_s^2\right] = \frac{10}{3} \qquad (5.3.8)$$

When $\sqrt{s} > 2m_b \sim 10$ GeV, $b\bar{b}$ can be produced, then

$$R = 3\left[Q_u^2 + Q_d^2 + Q_s^2 + Q_c^2 + Q_b^2\right] = \frac{11}{3} \qquad (5.3.9)$$

Because $2m_t \sim 350\,\mathrm{GeV}$, so

$$R = \begin{cases} 2 & \text{if } \sqrt{s} < 3\,\mathrm{GeV} \\ \frac{10}{3} & \text{if } 3\,\mathrm{GeV} < \sqrt{s} < 10\,\mathrm{GeV} \\ \frac{11}{3} & \text{if } 10\,\mathrm{GeV} < \sqrt{s} < 350\,\mathrm{GeV} \end{cases} \qquad (5.3.10)$$

The predictions in Eq. (5.3.10) agree with experimental measurements. The experimental results [7] are shown in Fig. 5.10. We can see from above that quarks have three different colors. This fact is verified by experiments. We shall discuss later that color symmetry $SU(3)_c$ is accurate local gauge symmetry. This symmetry will lead to the emerge of the quantum chromodynamics (QCD).

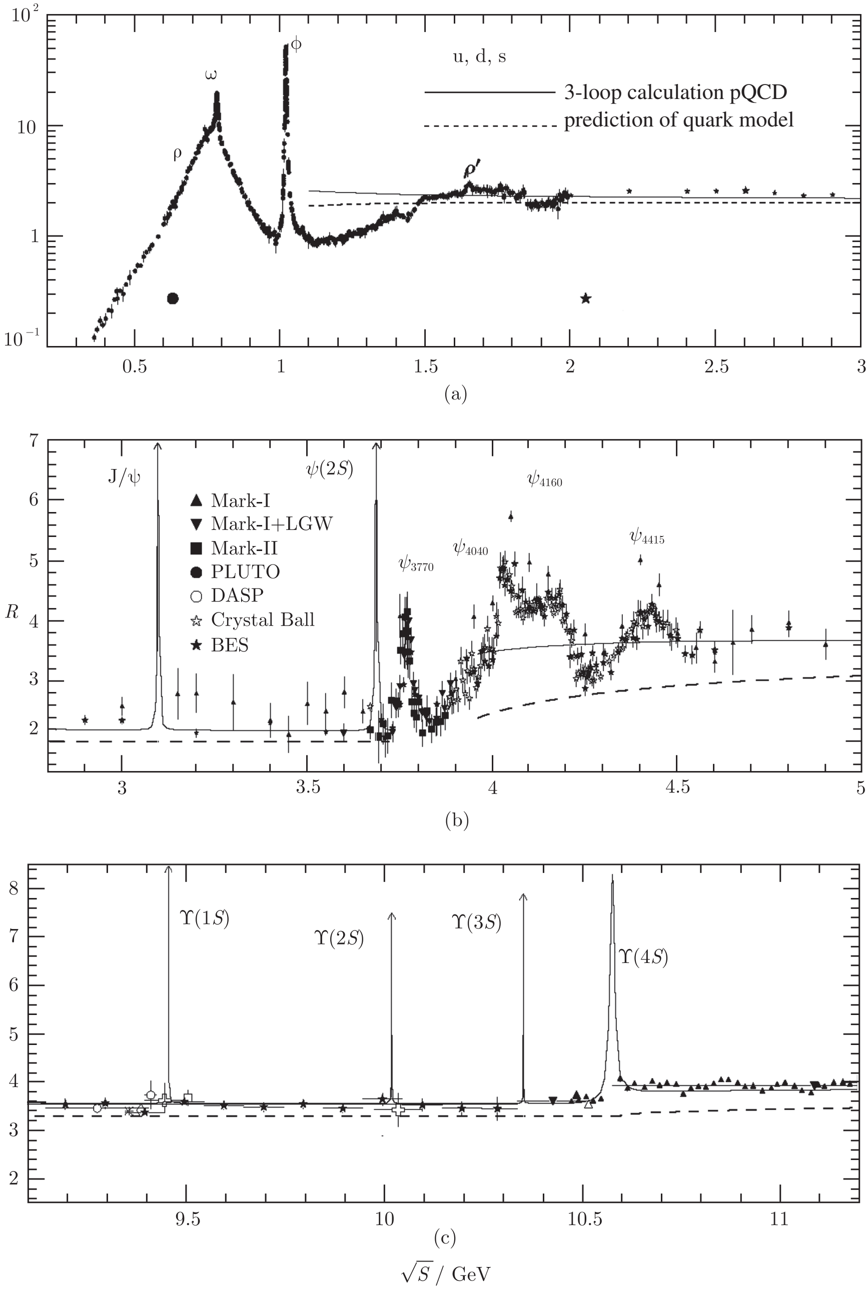


Fig. 5.10. The measured results of R.

5.4. Mass Formula of Hadrons

The success of the quark model shows that SU(3) flavor symmetry of strong interactions is a good symmetry. Later, we use $\mathrm{SU(3)_f}$ and $\mathrm{SU(3)_c}$ to denote flavor and color symmetries, respectively. We have mentioned in Sec. 5.3 that $\mathrm{SU(3)_c}$ is a accurate symmetry. If $\mathrm{SU(3)_f}$ is also a accurate symmetry, then the mesons or baryons in the same $\mathrm{SU(3)_f}$ multiplets (for instance $\underline{8}$ or $\underline{10}$) should have the same masses, respectively. But the experimental results are not so. Actually, we find in experiments that in the same $\mathrm{SU(3)_f}$ multiplet ($\underline{8}$ or $\underline{10}$) there is a several hundred MeV mass difference for different isospin multiplets. In Sec. 5.3, we already mention that for the fundamental basis $\underline{3}$ the masses of u, d, s quarks are also different. This means that $\mathrm{SU(3)_f}$ symmetry is not accurate, it must be broken.

The $\mathrm{SU(3)_f}$ mass formula is phenomenological; it is not accurately deduced from theory. Although the deduction methods are different, but essentially they are the same. Here we give a intuitive method. From the weight diagrams (Figs. 5.6 and 5.7), we can see that for baryon $\underline{10}$ multiplet the mass split is along Y-axis, and the gap of the mass difference for different isospin multiplets is equal. For example,

$$\begin{gathered} m_\Delta \sim 1232\ \mathrm{MeV},\ m_{\Sigma^*} \sim 1384\ \mathrm{MeV},\ m_{\Xi^*} \sim 1533\ \mathrm{MeV}, \\ m_{\Sigma^*} - m_\Delta = m_{\Xi^*} - m_{\Sigma^*} \sim 150\ \mathrm{MeV} \end{gathered} \tag{5.4.1}$$

So we can write for decuplet

$$m_{\underline{10}} = m_0 + aY \tag{5.4.2}$$

where m_0 and a are constants; their numerical values can be fixed by experiments. Equation (5.4.2) just satisfies the equal mass gap rule for $\left(\frac{3}{2}\right)^+$ baryon mass difference of adjacent supercharge multiplets shown in Eq. (5.4.1). So the mass formula in Eq. (5.4.2) is reasonable for $\left(\frac{3}{2}\right)^+$ baryon decuplet. But Eq. (5.4.2) is not suitable for baryon octet. The reason is as follows. Because $\Sigma^{\pm,0}$ is $I = 1$ isospin triplet, Λ is $I = 0$ singlet, their supercharge Y are all equal to zero, but the masses of them are different: $m_\Sigma \sim 1190$ MeV, $m_\Lambda \sim 1116$ MeV, the difference is 74 MeV, so for $\left(\frac{1}{2}\right)^+$ baryon octet we need to add isospin breaking term, for instance, the term like $\vec{I}^2 = I(I+1)$. But in order to make the mass formula also good for $\left(\frac{3}{2}\right)^+$ baryon decuplet, the added term should keep the equal mass gap rule for baryon decuplet. The term $\left[I(I+1) - \frac{1}{4}Y^2\right]$ is just a good choice. After adding this term, we can have a unified mass formula for both $\underline{10}$ and

$\underline{8}$ baryon multiplets

$$m = m_0 + aY + b\left[I(I+1) - \frac{1}{4}Y^2\right] \tag{5.4.3}$$

Equation (5.4.3) is called Gell-Mann–Okubo mass formula [9]. In Eq. (5.4.3), the three constants m_0, a, b are different for $\underline{10}$ and $\underline{8}$ multiplets. For baryon octet

$$\begin{aligned}
&\Lambda^0\colon I=0, Y=0; \quad \mathrm{p,n}\colon I=\frac{1}{2}, Y=1\\
&\Sigma^{\pm,0}\colon I=1, Y=0; \quad \Xi^0,\Xi^-\colon I=\frac{1}{2}, Y=-1
\end{aligned}$$

Substituting above parameters into Eq. (5.4.3) we get

$$\begin{aligned}
m_\Lambda &= m_0\\
m_\mathrm{N} &= m_0 + a + \frac{b}{2}\\
m_\Sigma &= m_0 + 2b\\
m_\Xi &= m_0 - a + \frac{b}{2}
\end{aligned} \tag{5.4.4}$$

Obviously from Eq. (5.4.4) we can deduce

$$3m_\Lambda + m_\Sigma = 2\left(m_\mathrm{N} + m_\Xi\right) \tag{5.4.5}$$

This is compatible with experiments. For baryon decuplet, Equation (5.4.3) gives

$$\begin{aligned}
m_\Delta &= m_0 + a + \frac{7b}{2}\\
m_{\Sigma^*} &= m_0 + 2b\\
m_{\Xi^*} &= m_0 - a + \frac{b}{2}\\
m_{\Omega^-} &= m_0 - 2a - b
\end{aligned} \tag{5.4.6}$$

Substituting m_Δ, m_{Σ^*}, m_{Ξ^*} into Eq. (5.4.6), we can fix the three constants a, b, m_0, then we get $m_{\Omega^-} \sim 1685$ MeV which is compatible with measured mass of Ω^-.

We can extend Eq. (5.4.3) directly to mesons. For the pseudoscalar meson octet, from Eq. (5.4.3) we have

$$m_\mathrm{K} = \begin{cases} m_0 + a + \dfrac{b}{2} \text{ for } Y=1 \text{ there are } \mathrm{K}^0, \mathrm{K}^+\\[2ex] m_0 - a + \dfrac{b}{2} \text{ for } Y=-1 \text{ there are } \mathrm{K}^-, \bar{\mathrm{K}}^0 \end{cases} \tag{5.4.7}$$

From Eq. (5.4.7) we see that $a = 0$ is the only right choice, so we have

$$\begin{aligned} m_{\mathrm{K}} &= m_0 + \frac{b}{2} \\ m_\pi &= m_0 + 2b \\ m_{\eta_8} &= m_0 \end{aligned} \tag{5.4.8}$$

From Eq. (5.4.8) we can derive

$$4m_{\mathrm{K}} = 3m_{\eta_8} + m_\pi \tag{5.4.9}$$

If we think that there is not much difference for η_8 and η, i.e. $m_{\eta_8} \approx m_\eta$, then Eq. (5.4.9) is not good. But if we use mass squared (matrix mesons obey Klein–Gordon equation in which mass squared matrix appears) in the mass formula, the result becomes better. So for mesons we use

$$m^2 = m_0^2 + aY + b\left[I(I+1) - \frac{Y^2}{4}\right] \tag{5.4.10}$$

Equation (5.4.8) becomes now

$$\left.\begin{aligned} m_{\mathrm{K}}^2 &= m_0^2 + \frac{b}{2} \\ m_\pi^2 &= m_0^2 + 2b \\ m_{\eta_8}^2 &= m_0^2 \end{aligned}\right. \Rightarrow 4m_{\mathrm{K}}^2 = 3m_{\eta_8}^2 + m_\pi^2 \tag{5.4.11}$$

If we take $m_{\eta_8} \approx m_\eta$, then Eq. (5.4.11) is consistent with experiments.

5.5. Mixing of Meson Singlet and Octet

In Secs. 5.2.1 and 5.2.2, we have already mentioned that there is mixing not only for η_1 and η_8, but also for ω_1 and ω_8. In fact, if SU(3)$_{\mathrm{f}}$ symmetry is accurate, the above mixing cannot happen. But because SU(3)$_{\mathrm{f}}$ symmetry is weakly broken, and in addition, $\eta_1 - \eta_8$, $\omega_1 - \omega_8$, all have $I = 0$, $Y = 0$, i.e. have same quantum numbers, so the breakdown of SU(3)$_{\mathrm{f}}$ symmetry induces mixing for $\eta_1 - \eta_8$, and $\omega_1 - \omega_8$. We first look at $\eta_1 - \eta_8$ mixing.

From Eqs. (5.2.14) and (5.2.15), we see that after mixing η, η' are experimentally observed mesons which are mass eigenstates, that is, the mass matrix (the matrix of the hamiltonian H) is diagonal when taking η, η' as basis.

Put $\zeta_1 \equiv \eta$, $\zeta_2 \equiv \eta'$, $\xi_1 \equiv \eta_8$, $\xi_2 \equiv \eta_1$, according to Eqs. (5.2.14) and (5.2.15), we have

$$\begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix} = U \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \tag{5.5.1}$$

that is

$$|\zeta_i\rangle = U_{ij}\,|\xi_j\rangle \tag{5.5.2}$$

where

$$U = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{5.5.3}$$

is a unitary and real matrix, thus

$$U^{+}U = U^{\mathrm{T}}U = 1 \tag{5.5.4}$$

When we take the physical eigenstates η, η' as basis, the Hamiltonian matrix (for particles at rest this is just mass matrix) is diagonal. So (note, U matrix is real)

$$\begin{aligned} \langle\zeta_i|\,H\,|\zeta_j\rangle &= U_{ik}^{*}U_{jl}\,\langle\xi_k|\,H\,|\xi_l\rangle \\ &= U_{ik}U_{jl}H_{kl} = U_{ik}H_{kl}\left(U^{\mathrm{T}}\right)_{lj} \\ &= \left(UHU^{\mathrm{T}}\right)_{ij} = m_i\delta_{ij} \end{aligned} \tag{5.5.5}$$

At the last step, the condition for particle at rest is used.

So we see that the mixing matrix U in Eq. (5.5.3) can diagonalize the mass matrix and certainly mass squared matrix, i.e.

$$UH^2U^{\mathrm{T}} = UHU^{\mathrm{T}}UHU^{\mathrm{T}} = \begin{pmatrix} m_\eta^2 & 0 \\ 0 & m_{\eta'}^2 \end{pmatrix} \tag{5.5.6}$$

At the basis of $\xi_1 \equiv \eta_8$, $\xi_2 \equiv \eta_1$, M^2 is not diagonal and has the following form:

$$M^2 = \begin{pmatrix} m_8^2 & m_{81}^2 \\ m_{18}^2 & m_1^2 \end{pmatrix} = U^{\mathrm{T}} \begin{pmatrix} m_\eta^2 & 0 \\ 0 & m_{\eta'}^2 \end{pmatrix} U \tag{5.5.7}$$

Assume M^2 is a real matrix, because of its Hermitian property, $M^{2+} = M^2$, we have

$$m_{81}^2 = m_{18}^2 \tag{5.5.8}$$

Substituting the matrix U into Eq. (5.5.7), we get

$$\begin{pmatrix} m_8^2 & m_{18}^2 \\ m_{18}^2 & m_1^2 \end{pmatrix} = \begin{pmatrix} m_\eta^2 \cos^2\theta + m_{\eta'}^2 \sin^2\theta & (m_{\eta'}^2 - m_\eta^2)\sin\theta\cos\theta \\ (m_{\eta'}^2 - m_\eta^2)\sin\theta\cos\theta & m_\eta^2 \sin^2\theta + m_{\eta'}^2\cos^2\theta \end{pmatrix} \tag{5.5.9}$$

Compare two sides of Eq. (5.5.9) and get

$$\begin{aligned} m_8^2 &= m_\eta^2 \cos^2\theta + m_{\eta'}^2 \sin^2\theta \\ m_1^2 &= m_\eta^2 \sin^2\theta + m_{\eta'}^2 \cos^2\theta \\ m_{18}^2 &= \left(m_{\eta'}^2 - m_\eta^2\right)\sin\theta\cos\theta \end{aligned} \tag{5.5.10}$$

From Eq. (5.5.10) we get

$$\begin{aligned} m_1^2 + m_8^2 &= m_{\eta'}^2 + m_\eta^2 \\ m_1^2 m_8^2 - m_{18}^4 &= m_\eta^2 m_{\eta'}^2 \\ \cos^2\theta &= (m_{\eta'}^2 - m_8^2)/(m_{\eta'}^2 - m_\eta^2) \end{aligned} \tag{5.5.11}$$

Dividing the two sides of the last formula of Eq. (5.5.10) by $\cos^2\theta$ and using the last formula of Eq. (5.5.11) we obtain

$$\tan\theta = \frac{m_{18}^2}{m_{\eta'}^2 - m_8^2} \tag{5.5.12}$$

From Eqs. (5.4.11), (5.5.11) and (5.5.10), we have

$$m_8^2 \equiv m_{\eta_8}^2 = \frac{4m_\mathrm{K}^2 - m_\pi^2}{3} \tag{5.5.13}$$

$$m_1^2 \equiv m_{\eta'}^2 + m_\eta^2 - m_8^2 \tag{5.5.14}$$

$$m_{18}^4 \equiv -m_\eta^2 m_{\eta'}^2 + m_1^2 m_8^2 \tag{5.5.15}$$

We can compute out m_{18}^2 and m_8^2 from Eq. (5.5.13) to Eq. (5.5.15), then calculate the value of $\tan\theta$ from Eq. (5.5.12).

Before we give the value of $\tan\theta$, we first look at the sign of m_{18}^2:

$$\begin{aligned} m_{18}^2 &= \langle\eta_1|\, H^2 \,|\eta_8\rangle \\ \eta_1 &= \frac{1}{\sqrt{3}}(\mathrm{u\bar{u}} + \mathrm{d\bar{d}} + \mathrm{s\bar{s}}) \\ \eta_8 &= \frac{1}{\sqrt{6}}(\mathrm{u\bar{u}} + \mathrm{d\bar{d}} - 2\mathrm{s\bar{s}}) \end{aligned} \tag{5.5.16}$$

Because the contribution of $\mathrm{u\bar{u}}$, $\mathrm{d\bar{d}}$ to the mass of η_1, η_8 is much less than that of $\mathrm{s\bar{s}}$, as it can be seen from the quark structure of η_1, η_8 in Eq. (5.5.16), the main contribution to m_{18}^2 comes from $s\bar{s}$ term in η_1 and $-2s\bar{s}$ term in η_8, so $m_{18}^2 < 0$.

Taking

$$\begin{aligned} m_\pi &= 135.6\,\mathrm{MeV}, \quad m_{\mathrm{K}_0} = 497.6\,\mathrm{MeV} \\ m_\eta &= 547.8\,\mathrm{MeV}, \quad m_{\eta'} \ = 957.8\,\mathrm{MeV} \end{aligned} \tag{5.5.17}$$

we get

$$m_8^2 = 0.3241\,(\mathrm{GeV})^2, \quad m_1^2 = 0.8934\,(\mathrm{GeV})^2$$

Substituting the above results into Eq. (5.5.15), we have

$$\begin{aligned} m_{18}^2 &= -0.1197\,(\mathrm{GeV})^2 \\ \tan\theta &= -0.2018 \\ \theta &= -11.4^\circ \end{aligned} \tag{5.5.18}$$

We should point out that the Gell-Mann–Okubo mass formula of Eq. (5.5.13) is only good for the lowest-order $O\left(p^2\right)$ in chiral perturbation theory, if we keep terms of $O\left(p^4\right)$ — the mass formula needs to be corrected [10].

For the mixing of $\omega - \phi$, we can do it following the same procedure for $\eta - \eta'$ mixing. For m^2 mass formula, we only need to make the following substitution:

$$\begin{aligned} \eta_1 \to \omega_1 &= \frac{1}{\sqrt{3}}(\mathrm{u\bar{u}} + \mathrm{d\bar{d}} + \mathrm{s\bar{s}}) \\ \eta_8 \to \omega_8 &= \frac{1}{\sqrt{6}}(\mathrm{u\bar{u}} + \mathrm{d\bar{d}} - 2\mathrm{s\bar{s}}) \end{aligned} \tag{5.5.19}$$

and

$$\begin{aligned} \eta &\to \phi \\ \eta' &\to \omega \end{aligned}$$

that is,

$$\begin{aligned} \phi &= \omega_8 \cos\theta_\mathrm{v} - \omega_1 \sin\theta_\mathrm{v} \\ \omega &= \omega_8 \sin\theta_\mathrm{v} + \omega_1 \cos\theta_\mathrm{v} \end{aligned} \tag{5.5.20}$$

Equation (5.5.10) is changed to

$$\begin{aligned} m_8^2 &= m_\phi^2 \cos^2\theta_\mathrm{v} + m_\omega^2 \sin^2\theta_\mathrm{v} \\ m_1^2 &= m_\phi^2 \sin^2\theta_\mathrm{v} + m_\omega^2 \cos^2\theta_\mathrm{v} \\ m_{18}^2 &= (m_\omega^2 - m_\phi^2) \sin\theta_\mathrm{v} \cos\theta_\mathrm{v} \end{aligned} \tag{5.5.21}$$

Similarly

$$\tan\theta_{\rm v} = \frac{m_{18}^2}{m_\omega^2 - m_8^2} \tag{5.5.22}$$

$$m_8^2 = \frac{1}{3}\left(4m_{\rm K^*}^2 - m_\rho^2\right) \tag{5.5.23}$$

It is the same as 0^- meson. Here for vector meson we also have $m_{18}^2 < 0$. Let

$$\begin{aligned} &m_{\rm K}^* \approx 0.892 \text{ GeV}, \quad m_\rho \sim 0.77\,6 \text{ GeV} \\ &m_\omega \sim 0.783 \text{ GeV}, \quad m_\phi \sim 1.01\,9 \text{ GeV} \end{aligned}$$

Then

$$m_8^2 \sim 0.860\,\text{GeV}^2 \tag{5.5.24}$$

so $m_\omega^2 - m_8^2 < 0$. From Eq. (5.5.22), we see that

$$\tan\theta_{\rm v} > 0$$

From Eq. (5.5.21), we can deduce

$$\begin{aligned} &m_1^2 + m_8^2 = m_\phi^2 + m_\omega^2 \\ &m_1^2 m_8^2 - m_{18}^4 = m_\omega^2 m_\phi^2 \\ &\cos^2\theta_{\rm v} = \frac{m_\omega^2 - m_8^2}{m_\omega^2 - m_\phi^2} \end{aligned} \tag{5.5.25}$$

We get

$$m_1^2 = m_\omega^2 + m_\phi^2 - m_8^2 \approx 0.791\,\text{GeV}^2 \tag{5.5.26}$$

Finally, we have

$$\tan\theta_{\rm v} \approx 0.846, \quad \theta_{\rm v} \sim 40^\circ \tag{5.5.27}$$

For the case of linear mass formula, the readers can solve it by themselves; we will not discuss it here.

In Eq. (5.2.17), we give the case for ideal mixing, where we take $\sin\theta_{\rm v} = 1/\sqrt{3}$, $\cos\theta_{\rm v} = \sqrt{2/3}$, so $\theta_{\rm v} \sim 35.3^\circ$. In that case

$$\begin{aligned} &\phi = -{\rm s\bar{s}} \\ &\omega = \frac{1}{\sqrt{2}}\left({\rm u\bar{u}} + {\rm d\bar{d}}\right) \end{aligned}$$

This result has not much difference with the measured one of Eq. (5.5.27).

5.6. OZI Rule

The spin parity of the vector mesons ω and ϕ is given by $J^{PC} = 1^{--}$. They decay to 0^- mesons mainly through strong interactions. We list their decay property in Table 5.3.

From Table 5.3, we see that although $m_\phi > m_\omega$, but the width $\Gamma_\phi < \Gamma_\omega$. 3π final state is more serious, because $\Gamma\left(\omega \to \pi^+\pi^-\pi^0\right) \gg \Gamma\left(\phi \to \pi^+\pi^-\pi^0\right)$. This means that some mechanism suppresses the decay of $\phi \to 3\pi$. We call this mechanism as Okubo–Zweig–Iizuka rule [11], simply called OZI Rule. OZI rule can be understand from the quark diagram for ω, ϕ decay. From Fig. 5.11, we see that the quark lines for the initial and final states of the decay $\omega \to \pi^+\pi^-\pi^0$ are connected through, so the decay rate is large. While in Fig. 5.12, the quark lines for the initial and final states of the decay $\phi \to \pi^+\pi^-\pi^0$ are not connected through, so it is forbidden.

Table 5.3. The main decay modes and branching ratios of ω and ϕ.

	m (MeV)	Γ (MeV)	Mode	Branching ratio
ω	782.65 ± 0.12	8.49 ± 0.06	$\pi^+\pi^-\pi^0$	$(89.2 \pm 0.7)\%$
			$\pi^0\gamma$	$(8.92 \pm 0.24)\%$
			K^+K^-	$(49.2 \pm 0.6)\%$
ϕ	1019.455 ± 0.029	4.26 ± 0.24	$K_L K_S$	$(34.0 \pm 0.5)\%$
			$\pi^+\pi^-\pi^0 + \rho\gamma$	$(15.25 \pm 0.35)\%$

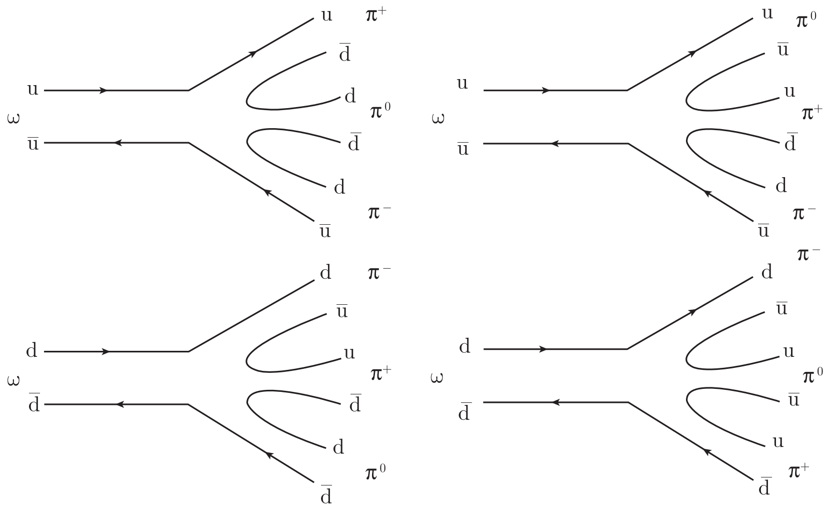

u π^+
$\bar{d}$
d π^0
$\bar{d}$
d π^-
$\bar{u}$
u
ω
$\bar{u}$
u π^0
$\bar{u}$
u π^+
$\bar{d}$
d π^-
$\bar{u}$
u
ω
$\bar{u}$
d π^-
$\bar{u}$
u π^+
$\bar{d}$
d π^0
$\bar{d}$
d
ω
$\bar{d}$
π^- d
$\bar{u}$
u π^0
$\bar{u}$
u π^+
$\bar{d}$
d
ω
$\bar{d}$

Fig. 5.11. Quark diagrams for the decay of $\omega \to \pi^+\pi^-\pi^0$.

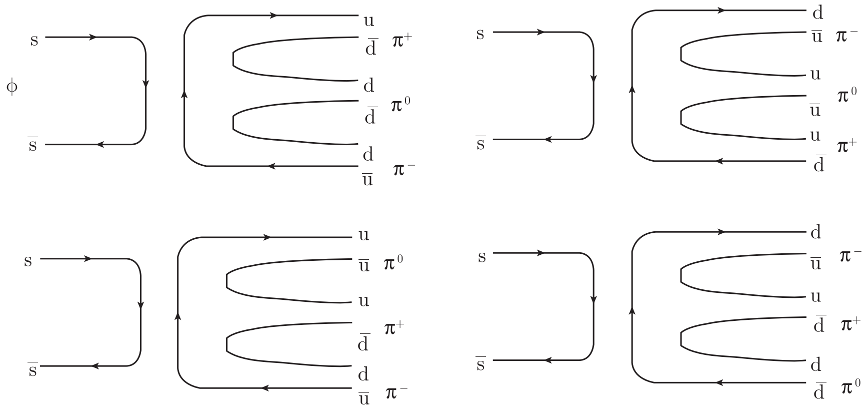

s
ϕ
$\bar{s}$
u
$\bar{d}$
π⁺
d
$\bar{d}$
π⁰
d
$\bar{u}$
π⁻
s
$\bar{s}$
d
$\bar{u}$
π⁻
u
$\bar{u}$
π⁰
u
$\bar{d}$
π⁺
s
$\bar{s}$
u
$\bar{u}$
π⁰
u
$\bar{d}$
π⁺
d
$\bar{u}$
π⁻
s
$\bar{s}$
d
$\bar{u}$
π⁻
u
$\bar{d}$
π⁺
d
$\bar{d}$
π⁰

Fig. 5.12. Quark diagram for the decay of $\phi \to \pi^+\pi^-\pi^0$.

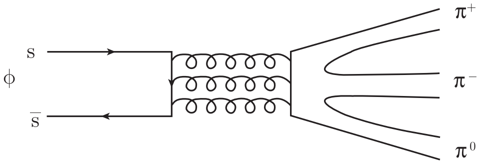

s
ϕ
$\bar{s}$
π⁺
π⁻
π⁰

Fig. 5.13. Decay diagram of $\phi \to \pi^+\pi^-\pi^0$ in QCD.

OZI can be phrased as: The decays in which the quark lines for initial and final states are not connected through are suppressed compare with that connected through.

We will see later that the decay widths of J/ψ ($c\bar{c}$ bound state) and Y ($b\bar{b}$ bound state) are all small. This can also be explained by OZI rule.

According to QCD (see Chapter 9), OZI suppressed processes are actually higher-order ones in perturbation theory. In Fig. 5.12, the decay of ϕ to 3π is actually higher-order process of exchanging three gluons between initial and final states. We show the decay diagram in Fig. 5.13.

Because the decay for spin 1 vector meson to two photons or two gluons is forbidden, the intermediate state in Fig. 5.13 must have three gluons.

5.7. SU(6) Symmetry

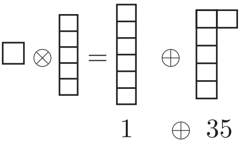
We can put in spin the $SU(3)_f$ quark model. Because the three flavored quarks u, d, s all have two states of spin up and down, so there are six states altogether. Using

$$q^i_\alpha = (\mathrm{u}_1, \mathrm{u}_2; \mathrm{d}_1, \mathrm{d}_2; \mathrm{s}_1, \mathrm{s}_2) \tag{5.7.1}$$

to express these six states, where $\alpha = 1, 2$ denote two states of $s_z = \frac{1}{2}, -\frac{1}{2}$, $i = 1, 2, \ldots, 6$ denote u_1, u_2, d_1, d_2, s_1, s_2. Taking q^i_α as the fundamental representation of hexaplet $\underline{6}$, we can construct the quark model of SU(6) symmetry.

For mesons ($\mathrm{q}\bar{\mathrm{q}}$ states)

$$\mathrm{q} \times \bar{\mathrm{q}} = \underline{6} \otimes \underline{6}^* =$$



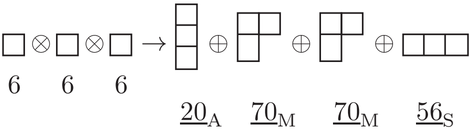
(5.7.2)

For decomposition of SU(6) to subgroup SU(3) × SU(2), we only give the result here. The details can be found in Itzykson's paper [6].

The decomposition of $\mathrm{SU}(6) \to \mathrm{SU}(3)_\mathrm{f} \times \mathrm{SU}(2)_\mathrm{spin}$ is

$$\begin{array}{lll}
\mathrm{SU}(6) & \to \mathrm{SU}(3)_\mathrm{f} \times \mathrm{SU}(2)_\mathrm{spin} & \\
\underline{1} & \to (1,1) & \text{fills } \eta_1 \ (0^-\text{meson singlet}) \\
\underline{35} & \to (8,3) \oplus (1,3) \oplus (8,1) & \\
 & \quad\ \downarrow \qquad\quad \downarrow \qquad\quad \downarrow & \\
 & (\rho, \mathrm{K}^*, \omega_8),\ \omega_1, (\mathrm{K}, \pi, \eta_8) &
\end{array}$$

for baryon qqq = 6 ⊗ 6 ⊗ 6



(5.7.3)

where the symbol A denotes totally antisymmetric, M denotes mix-symmetric, S denotes totally symmetric. If we use ϕ to denote $\mathrm{SU}(3)_\mathrm{f}$ wave function, and χ to denote SU(2) spin wave function, then, in the SU(6) wave function, $\phi_\mathrm{S}\chi_\mathrm{S}$, $\phi_\mathrm{A}\chi_\mathrm{A}$ are symmetric and $\phi_\mathrm{S}\chi_\mathrm{A}$, $\phi_\mathrm{A}\chi_\mathrm{S}$ are antisymmetric respectively. The decomposition of $\underline{56}_\mathrm{S}$ to $\mathrm{SU}(3)_\mathrm{f} \times \mathrm{SU}(3)_\mathrm{spin}$ is

$$\begin{array}{lccc}
\underline{56}_\mathrm{S} = & (10,4) & \oplus & (8,2) \\
 & \text{fill } \left(\dfrac{3}{2}\right)^+ \text{ baryons} & & \text{fill } \left(\dfrac{1}{2}\right)^+ \text{ baryons}
\end{array}$$

The two $\underline{70}_\mathrm{M}$ and one $\underline{20}_\mathrm{A}$ are not filled up at present, perhaps some orbital excitation states can fill in.

For the detail about the structure of SU(6) wave function, either the readers can deduce it by themselves or refer to [8].

5.8. Orbital Excitation States and Radial Excitation States, Multi-quark States and Exotic States

Already we only discussed mesons and baryons in ground states. There are some other cases we did not involve. These cases are the orbital excitation states, radial excitation states, multi-quark states, and exotic states as follows:

5.8.1. *Orbital and radial excitation states*

The quark structure of mesons is $q\bar{q}$. If L is the relative angular momentum of $q\bar{q}$, then the parity of the $q\bar{q}$ system is

$$P = (-1)^{L+1} \tag{5.8.1}$$

The reason is that the relative parity of quark and antiquark is (-1), the contribution of space wave function to parity is $(-1)^L$. If we introduce the charge conjugation parity (particle and antiparticle conjugation) C for $q\bar{q}$ system, then $q\bar{q}$ system can be viewed as identical particle system, so when exchanging q and $\bar{q}$ the total wave function should be antisymmetric, i.e.

$$C(-1)^L(-1)^{S+1} = -1 \tag{5.8.2}$$

so we get the charge conjugation parity

$$C = (-1)^{L+S} \tag{5.8.3}$$

The results of Eqs. (5.8.1) and (5.8.3) agree with that of rigorous deduction in Chapter 2.

Note, only when q and $\bar{q}$ have same flavor quantum number, then the $q\bar{q}$ system can have definite C parity.

The values of $J^{PC}(^{2S+1}L_J)$ for $q\bar{q}$ system are listed in Table 5.4.

Orbital and radial excitation states existed in atomic physics long time ago, for instance in e^+e^- bound states (positronium). We have already found its orbital and radial excitation states before.

In particle physics, the orbital and radial excitation states in $(c\bar{c})$, $(b\bar{b})$ systems have been found. For detail, please refer the later part of the book.

For three quark states qqq, the three quarks are all in s-wave for ground state. For the first excitation state one of the three quarks is in p wave,

Table 5.4. Values of $J^{PC}(^{2S+1}L_J)$ for $q\bar{q}$ system.

	$S = 0\ (CP = -1)$	$S = 1\ (CP = +1)$
$L = 0$	$0^{-+}(^1S_0)$	$1^{--}(^3S_1)$
$L = 1$	$1^{+-}(^1P_1)$	$0^{++}(^3P_0)$
		$1^{++}(^3P_1)$
		$2^{++}(^3P_2)$
$L = 2$	$2^{-+}(^1D_2)$	$1^{--}(^3D_1)$
		$2^{--}(^3D_2)$
		$3^{--}(^3D_3)$

Note: States of $J^{PC} = 0^{--}, 0^{+-}, 1^{-+}, 2^{+-}$ etc. do not exist in the above $q\bar{q}$ system, these states are called exotic states.

then the space wave function of the three quarks is of mixed symmetry, its SU(6) wave function is also in mixed symmetry. For the decomposition of $6 \otimes 6 \otimes 6$ in SU(6), the representation $\underline{70}_{\mathrm{M}}$ (see Sec. 5.7.3) belongs to mixed symmetry. The decomposition of $\underline{70}_{\mathrm{M}}$ into $\mathrm{SU}(3)_{\mathrm{f}} \times \mathrm{SU}(2)_{\mathrm{spin}}$ is

$$\begin{array}{ccccccc}
\underline{70}_{\mathrm{M}} \to (10,2) & + & (8,2) & + & (8,4) & + & (1,2) \\
\downarrow & & \downarrow & & \downarrow & & \downarrow \\
\text{fill}\left(\frac{1}{2}\right)^-, \left(\frac{3}{2}\right)^- & & \text{fill}\left(\frac{1}{2}\right)^-, \left(\frac{3}{2}\right)^- & & \text{fill}\left(\frac{1}{2}\right)^-, \left(\frac{3}{2}\right)^-, \left(\frac{5}{2}\right)^- & & \text{fill}\left(\frac{1}{2}\right)^-, \left(\frac{3}{2}\right)^-
\end{array}$$

The investigation for the fill up of the qqq excitation states is in progress.

5.8.2. *Multi-quark states and exotic states*

We concern only four quark states $(q\bar{q})(q\bar{q})$ and five quark states $qqq\,(q\bar{q})$. Of course, there are states consist of more quarks and antiquarks. At present, there are already candidates of four quark states in experiments, for example, at BEPCII (BESIII) a bench of new resonances are observed, where there are $Z_c^{\pm}(3900)$[12], $Z_c^{\pm}(4200)$[13], $Z_c^{\pm}(3885)$[14] and $Z_c^{\pm}(4025)$[15]. The four resonances are all possibly four quark states of $cu\bar{c}\bar{d}$ or $cd\bar{c}\bar{u}$. It certainly needs further theoretical and experimental investigation.

In addition, there are also exotic states in glueballs and hybrids. More can be found in Chapter 9.

5.9. The Discovery of c, b, t Quarks

The existence of c quark (charm) is verified in 1974 by Ting's group at Brookhaven National Laboratory (BNL) and Richter's group at Stanford Linear Accelerator Center (SLAC) simultaneously.

In the experiment of BNL [16], they use proton to bombard beryllium target and observe the invariant mass of e^+e^- pair, i.e.

$$\begin{aligned} &\mathrm{p} + \mathrm{Be} \to \mathrm{J}/\psi + \text{anything} \\ &\qquad\qquad\quad \hookrightarrow \mathrm{e}^+\mathrm{e}^- \end{aligned} \tag{5.9.1}$$

while in the experiment of SLAC [17], they observe the following process:

$$\mathrm{e}^+\mathrm{e}^- \to \psi \to \mathrm{e}^+\mathrm{e}^-, \mu^+\mu^- \tag{5.9.2}$$

The new particle discovered in experiments J/ψ has $J^{PC} = 1^{--}$, $M = (3097 \pm 1)\,\mathrm{MeV}$, $\Gamma = (69 \pm 7)\,\mathrm{keV}$. So its decay width is very small, that is, its lifetime is extraordinarily long (compare with its large mass). Actually J/ψ is just $c\bar{c}$ bound state.

Later, $\mathrm{D}^\pm$ and D^0 mesons are observed. For quark structure $\mathrm{D}^+ = \mathrm{c}\bar{\mathrm{d}}$, $\mathrm{D}^- = \bar{\mathrm{c}}\mathrm{d}$, $\mathrm{D}^0 = \mathrm{c}\bar{\mathrm{u}}$. The electric charge of c quark is 2/3 of the proton charge, its isospin is zero. We will come back to this later

In 1977, L. Lederman's group [18] discovered two new long lifetime particles Y (9.5 GeV) and Y′(10 GeV) at Fermi Laboratory proton synchrotron, where

$$m_{\mathrm{Y}} = (9.46 \pm 0.01)\,\mathrm{GeV}, \quad \Gamma_{\mathrm{Y}} = (42 \pm 15)\,\mathrm{keV}$$
$$m_{\mathrm{Y}'} = (10.02 \pm 0.01)\,\mathrm{GeV}, \quad \Gamma_{\mathrm{Y}'} = (30 \pm 10)\,\mathrm{keV}$$

Here Y′ is the radial excitation state of Y and Y, Y′ are all $(\mathrm{b}\bar{\mathrm{b}})$ bound states. The electric charge of b quark is $-\frac{1}{3}$ proton charge, and its isospin is zero. The discovery of Y and Y′ verifies the existence of the b quark. The English name of the b quark is bottom or beauty, its discovery process is

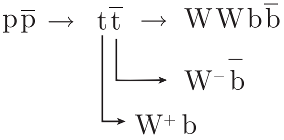
$$\mathrm{p} + \mathrm{Be}, \mathrm{Cu}, \mathrm{Pt} \to \mu^+\mu^- + \text{anything}$$

In fact, it is

$$\begin{aligned} &\mathrm{p} + \mathrm{Be},\ \mathrm{Cu},\ \mathrm{Pt} \to \mathrm{Y}(\mathrm{Y}') + \text{anything} \\ &\qquad\qquad\qquad\qquad\ \ \hookrightarrow \mu^+\mu^- \end{aligned}$$

In 1994, the top quark t (top) [19] is discovered at Fermi Laboratory. CDF Collaboration Group observes the following process at Fermi Laboratory.

$\mathrm{p}-\bar{\mathrm{p}}$ collider (Tevatron),



$$\begin{tikzcd} \mathrm{p\bar{p}} \arrow[r] & \mathrm{t\bar{t}} \arrow[r] \arrow[dr] \arrow[ddr] & \mathrm{WWb\bar{b}} \\ & & \mathrm{W^{-}\bar{b}} \\ & & \mathrm{W^{+}b} \end{tikzcd}$$

They discovered the top quark t and measured the top quark mass as about 176 GeV. Because the t quark is very heavy, and its lifetime is extremely shot, it decays immediately after its creation and no time to form any hadronic bound state. That is why we have not observed any $(\mathrm{t\bar{t}})$ bound states and $(\mathrm{t\bar{q}})$ bound states. The electric charge of the t quark is 2/3 proton charge, its isospin $I = 0$.

Until now, the discovered quarks are totally six flavors, each flavor has three different colors, so we have

$$\mathrm{u, d, s, c, b, t},$$

plus three colors we have total 18 different quarks. Because c, b, t quark are all heavy, $m_{\mathrm{c}} \sim 1.27\,\mathrm{GeV}$, $m_{\mathrm{b}} \sim 4.29\,\mathrm{GeV}$, $m_{\mathrm{t}} \sim 176\,\mathrm{GeV}$, so c, b, t are also called heavy flavors. But only the flavor symmetry for u, d, s, i.e. the $\mathrm{SU(3)_f}$, is good, although it still has very small breakdown. As the mass difference for c, b, t and u, d, s is very large, the $\mathrm{SU(4)_f}$ flavor symmetry including c quark, the $\mathrm{SU(5)_f}$ flavor symmetry including b quark, and the $\mathrm{SU(6)_f}$ flavor symmetry including t quark are all not meaningful. That is why, we only discuss $\mathrm{SU(3)_f}$ flavor symmetry.

5.10. Quark Confinement

Scientists have been looking for free quarks since 1964, but until now no matter from high-energy accelerator experiments or cosmic-ray experiments they have not found any free quarks. People tried to look for free quarks with color charge and fractional electric charge in sea, mineral substances, even at the moon, but no success. It is worthwhile to mention that there is one experiment which reported the existence of free particles with fractional charge [20]. In this experiment, they levitate small diamagnetic niobium spheres (100 μg) in a oscillated magnetic field (similar to the Millkan's oil-drop device). They made 39 measurements in which 5 measurements had a negative charge of mean value -0.343 ± 0.001, and 9 gave a positive charge of mean value 0.328 ± 0.007. The left 25 measurements showed an integral charge of mean value 0.001 ± 0.003. If we include the total number of nucleons in the niobium sphere the number of fractional charges observed corresponds to about 10^{-20} quarks per nucleon of niobium.

The other experiment utilized ferromagnetic cylinders (200μg) suspended in a magnetic field [21]. As a result only integral charges were found. This corresponds to the probability for finding free quarks in iron is less than 3×10^{-21} quark/per nucleon.

We can see, from the above, that nucleon consists of three valence quarks (plus an infinite number of (q$\bar{\text{q}}$) pair of sea quarks and gluons), but we have never observed free quarks. This phenomenon is called "quark confinement". Until now we have never observed free gluons (it carries color charge), so some times we call these two phenomena as "color confinement".

The explanation for quark confinement is a difficult issue of particle physics today. There is no commonly recognized theory at present. But there are many theories including phenomenological ones which try to explain quark confinement. One of these, for example, is the Cornell Potential model [22] which tries to explain the spectra of the charmonium and bottomonium

$$V(r) = -\frac{k}{r} + \frac{r}{a^2} \tag{5.10.1}$$

It includes the term of the linear potential r/a^2. This term makes it impossible to separate c and $\bar{c}$ or b and $\bar{b}$, because it needs to do infinite work.

Another example is T.D. Lee's work [23]. He assumes that the QCD vacuum is a kind of media with antiscreening property (compare to dielectric media), so explains quark confinement.

Because quark confinement potential has the characteristics of long-range force, the Monte Carlo method of Lattice Gauge Theory is more promising for solving this problem. At present, the international physics community pays more attention to that direction.

For a review, please see the review article [24].

References

1. Fermi E, Yang C N. *Phys. Rev.*, 1949, 76: 1739.
2. Sakata S. *Prog. Theor. Phys.*, 1956, 16: 686.
3. Gell-Mann M. Caltech Report CTSL-20, 1961; Gell-Mann M. *Phys. Rev. Lett.*, 1962, 125: 1067; Ne'eman *Y. Nucl. Phys.*, 1961, 26: 222.
4. Barnes V E, *et al. Phys. Rev. Lett.*, 1964, 12: 204.
5. Zweig G. CERN Report; Gell-Mann M. *Phys. Rev. Lett.*, 1964, 8: 214.
6. Itzykson C. *Rev. Mod. Phys.*, 1966, V38: 95.
7. Particle Data Group. *J. Phys. G: Nucl. Part. Phys.*, 2010, V37: 075021.

8. Close F E. *An Introduction to Quarks and Partons.* Academic Press, 1979: 46.
9. Okubo S. *Prog. Theor. Phys.*, 1962, 27: 949; Gell-Mann M, Oaks R J, Renner B. *Phys. Rev.*, 1968, 175: 2195.
10. Venugopal E P, Holstein B R. *Phys. Rev.*, 1998, D57: 4397.
11. Okubo S. *Phys. Lett.*, 1963, 5: 165; Zweig G. *Symmetries in Elementary Particle Physics.* ed. A. Zichichi, Academic Press, 1965; Iizuka *J. Prog. Theor. Phys. Suppl.*, 1966, 37–38: 21.
12. Ablikim M, *et al.* (BESIII Collab.). *Phys. Rev. Lett.*, 2013, 110: 252001.
13. Ablikim M, *et al.* (BESIII Collab.). *Phys. Rev. Lett.*, 2013, 111: 242001.
14. Ablikim M, *et al.* (BESIII Collab.). *Phys. Rev. Lett.*, 2014, 112: 022001.
15. Ablikim M, *et al.* (BESIII Collab.). arXiv: 2014, 1308: 2760.
16. Aubert J J, *et al. Phys. Rev. Lett.*, 1974, 33: 1404.
17. Augustin J E, *et al. Phys. Rev. Lett.*, 1974, 33: 1406.
18. Herb S W, *et al. Phys. Rev. Lett.*, 1977, 39: 252; Innes W R, *et al. Phys. Rev. Lett.*, 1977, 39: 1240.
19. Abe F, *et al.* (CDF Collab.). *Phys. Rev. Lett.*, 1995, 74: 2626.
20. La Rue G S, *et al. Phys. Rev. Lett.*, 1977, 38: 1011: 981, 46: 967.
21. Gallinaro G, *et al. Phys. Rev. Lett.*, 1977, 38: 1255.
22. Eichten E, *et al. Phys. Rev.*, 1978, D17: 3090.
23. Lee T. D., *Particle Physics and Introduction to Field Theory.* Science Press, 1980.
24. Alkofer R, Greensite J. arXiv: hep-ph/0610365; ’t Hooft, G. *Nucl. Phys.*, 2003, A721: 3–19.

Chapter 6

Electromagnetic Interaction

Among three interactions — strong interaction, electromagnetic interaction and weak interaction — that govern the microcosm, electromagnetic interaction is the most clearly understood interaction up to now. Far more has been known about electromagnetic interaction than about any other interactions. This is due to the fact that the electromagnetic coupling constant, the fine structure constant α, is a small quantity: the fine structure constant $\alpha \approx 1/137$, that is a small enough number. This allows the theoretical calculation of any electromagnetic interaction to be expanded in series of α. Usually the theoretical calculation can be precise enough by just taking the first few terms of this expansion. So far, all experimental tests of electromagnetic interactions with sufficient precision are consistent with quantum electrodynamics (QED), that is the theory that describes microscopic electromagnetic interactions. This theory was developed by R. Feynman, J. Schwinger and S. Tomonaga in 1940–1950s [1–9]. They developed renormalization methods, which can be used to systematically eliminate the infinity in the calculation of higher-order expansions. The infinity in the higher-order calculation is absorbed into the electric charge, mass and other parameters of the charged particles. These parameters containing the infinity are then substituted by physically measured quantities to eliminate the infinity. Renormalization makes the calculation of higher-order correction of QED meaningful. The calculation of higher-order corrections not only improve the accuracy of theoretical calculation, but also lead to direct observation effect in physics. The explanations of some experimental measurements, such as Lamb shift [10], anomalous magnetic moments of electron and muon [11, 12], etc., were all realized by the calculation of higher-order terms in QED [6, 13–15].

QED is a theory of describing interactions between leptons and electromagnetic field, where leptons are viewed as point particles. The experimental test of QED has been carried out to as deep as $10^{-16} - 10^{-19}$m. No deviation between experiment and theory has been found at such a microscopic scale [16–19]. This indicates that no inner structure of electron has been found at such microscopic scale. Because of this, leptons such as electrons are often used to probe structure of other particles. From the mid-1950s to the late 1960s, a large number of experiments using high speed electrons to bombard nucleons were performed at Stanford Linear Accelerator Center (SLAC) in the United States. The measurements of the electron scattering cross section and the electron angular distribution showed that there is indeed inner structure in nucleons [20–23]. Nucleon cannot be viewed as point particle. There are specific distributions for both the electric charge and magnetic moment for nucleon [20]. And the successive experiment of electron–nucleon deep inelastic scattering found that, as the scattering angle increases, the scattering cross section showed great similarity to that of electron-electron scattering [21–23]. Inspired by this result, in 1968, J.D. Bjorken predicted the scaling phenomenon using current-algebra method [24]. Soon later Feynman proposed the parton model for nucleon [25]. And by calculation, it was proved that the parton model can derive scaling phenomenon, which is the important support to the parton model. Parton model is actually consistent with quark model. Partons include not only valence quarks in nucleon, but also sea quarks.

The study of electromagnetic interaction not only helps people to understand the interaction between electron and electromagnetic field, but also helps people to understand the microscopic laws of the generation and transition of electron, positron and photons. It also helps people to better understand the structure of the nucleons. At first a brief introduction to QED will be given in this chapter, and then some scattering processes for testing QED will be discussed successively. Two basic processes are as follows: (1) Møller scattering $e^-e^- \to e^-e^-$; (2) Bhabha scattering $e^+e^- \to e^+e^-$. The elastic and inelastic scattering processes of electron and nucleon are discussed. Finally, the parton model for nucleon is introduced.

6.1. QED and Its Feynman Rule

QED is a quantum theory of the interaction between electrons or other charged leptons and electromagnetic field. It can describe not only the creation and annihilation of electron and positron, but also the emission and absorption of photons. The interaction between electron and electromagnetic field in QED can be introduced either by analogy with classical electrodynamics or directly based on the principle of gauge invariance.

In classical electrodynamics, the interaction between charged particles with a charge of Qe and the electromagnetic field shows up as the Lorentz force exerted by the electromagnetic field on the charged particles $\vec{F}_B = Qe\vec{v} \times \vec{B}$ and the electric field force $\vec{F}_E = Qe\vec{E}$, where $\vec{B}$ and $\vec{E}$ are the magnetic-flux density and electric field strength, respectively. Define $\vec{A}$ and φ are the vector and scalar potentials, respectively. The relations between the potential and the electric field $\vec{E}$ and magnetic field $\vec{B}$ are

$$\begin{gathered} \vec{E} = -\nabla\varphi - \frac{\partial \vec{A}}{\partial t} \\ \vec{B} = \nabla \times \vec{A} \end{gathered} \tag{6.1.1}$$

Replace the momentum $\vec{p}$ in the classical Hamiltonian for free particle with $\vec{p} - Qe\vec{A}$, then the classical Hamiltonian becomes

$$H = \sqrt{\left(\vec{p} - Qe\vec{A}\right)^2 + m^2} + Qe\varphi \tag{6.1.2}$$

where $\vec{p}$ is the canonical momentum. Electromagnetic field is contained in the above Hamiltonian, and the Hamiltonian satisfies Hamilton's equation

$$\begin{gathered} \frac{\mathrm{d}q_i}{\mathrm{d}t} = \frac{\partial H}{\partial p_i} \\ \frac{\mathrm{d}p_i}{\mathrm{d}t} = -\frac{\partial H}{\partial q_i} \end{gathered} \tag{6.1.3}$$

where q_i is the generalized coordinate, and p_i the canonical momentum. From this Hamiltonian, which contains the electromagnetic field, and considering the Hamiltonian equation that it satisfies, the electric force and Lorentz force on the charged particles in an electromagnetic field can be derived correctly

$$\vec{F} = Qe(\vec{E} + \vec{v} \times \vec{B}) \tag{6.1.4}$$

Therefore, Eq. (6.1.2) is the classical Hamiltonian that correctly reflects the interaction between electromagnetic field and charged particles. Equation (6.1.2) also tells us that in order to consider the interaction between charged particle and electromagnetic field, it is only necessary to make the replacement $\vec{p} \to \vec{p} - Qe\vec{A}$ and $H \to H - Qe\varphi$ in the Hamiltonian of free particles.

Replace the momentum and energy in the Hamiltonian in Eq. (6.1.2) with quantum mechanical operators

$$\begin{aligned} \vec{p} - Qe\vec{A} &\to -\mathrm{i}\nabla - Qe\vec{A} \\ H - Qe\varphi &\to \mathrm{i}\frac{\partial}{\partial t} - Qe\varphi \end{aligned} \tag{6.1.5}$$

and introduce four-vector potential for electromagnetic field $A^\mu = (\varphi, \vec{A})$, then the two equations in Eq. (6.1.5) can be written in a unified manner

$$\mathrm{i}\partial^\mu - QeA^\mu = \mathrm{i}\left(\partial^\mu + \mathrm{i}QeA^\mu\right) \tag{6.1.6}$$

Equation (6.1.6) indicates that in order to consider the interaction between charged particle and electromagnetic field in quantum theory, we just need to do the replacement

$$\partial_\mu \to \partial_\mu + \mathrm{i}QeA_\mu \tag{6.1.7}$$

in the Hamiltonian for free particle. Therefore do the above replacement in the free-particle Lagrangian

$$\mathcal{L}_{\text{free}} = \bar{\psi}\left(\mathrm{i}\gamma^\mu\partial_\mu - m\right)\psi \tag{6.1.8}$$

we get

$$\mathcal{L} = \bar{\psi}\mathrm{i}\gamma^\mu\left(\partial_\mu + \mathrm{i}QeA_\mu\right)\psi - m\bar{\psi}\psi \tag{6.1.9}$$

This is the Lagrangian of QED that describes the interaction between electron and electromagnetic field.

In Eq. (6.1.9), the form of QED Lagrangian is also the form required by the local gauge invariance. The absolute phase of a fermion field is not a directly measurable quantity. A theory that correctly reflects the dynamics of fermions should be formally independent on such phases. Perform a global phase transformation for the fermion field in Eq. (6.1.9), that is,

$$\psi \to \psi' = \mathrm{e}^{-\mathrm{i}Qe\theta}\psi \tag{6.1.10}$$

where θ is an arbitrary phase angle. It can be verified that Eq. (6.1.9) is indeed invariant under such a transformation, that is, there is global gauge invariance for QED.

In addition, the four-vector A^μ of electromagnetic field is not direct observables either. The physical observables are the electric field $\vec{E}$ and magnetic flux density $\vec{B}$. That is, the corresponding four-vector potential

A^μ relevant to definite electric field $\vec{E}$ and magnetic field $\vec{B}$ can take large amount of possible values, which can be verified with Eq. (6.1.1).

Pick any scalar function $\theta(x)$ that depends on the coordinates of space and time, and transform φ and $\vec{A}$ as follows:

$$\begin{aligned} \vec{A} \to \vec{A}' &= \vec{A} - \nabla\theta(x) \\ \varphi \to \varphi' &= \varphi + \frac{\partial}{\partial t}\theta(x) \end{aligned} \tag{6.1.11}$$

The transformation above can also be written more concisely as

$$A_\mu \to A'_\mu = A_\mu + \partial_\mu \theta(x) \tag{6.1.12}$$

Such transformation is called gauge transformation of electromagnetic field. Under such transformation, it can be verified from Eq. (6.1.1) that $\vec{E}$ and $\vec{B}$ are all invariant. This indicates that there are quite large amount of degrees of freedom for vector A^μ in physics. All A_μ's that are related by the transformation of Eq. (6.1.12) correspond to the same physical quantity. Therefore, a theory that correctly describes the interaction of electron and electromagnetic field must be invariant under gauge transformation given in Eq. (6.1.12). It can be easily seen from Eq. (6.1.9) that, if transform the electromagnetic field under gauge transformation in a simple manner, Lagrangian $\mathcal{L}$ is not invariant.

The method to solve this contradiction is to consider gauge transformation of electromagnetic and fermion field jointly. At first, it is necessary to extend the globe gauge transformation in Eq. (6.1.10) to local gauge transformation, that is, to extend the constant θ in Eq. (6.1.10) to a function $\theta(x)$ that depends on space–time coordinate. At the same time view the $\theta(x)$ in Eq. (6.1.10) and that in Eq. (6.1.12) as the same quantity. As electromagnetic field is transformed under gauge transformation of Eq. (6.1.12), fermion field is also transformed under the following gauge transformation:

$$\psi \to \psi' = \mathrm{e}^{-\mathrm{i}Qe\theta(x)}\psi \tag{6.1.13}$$

Under the joint transformation of Eqs. (6.1.12) and (6.1.13), we can obtain

$$\left(\partial_\mu + \mathrm{i}QeA'_\mu\right)\psi' = \mathrm{e}^{-\mathrm{i}Qe\theta(x)}\left(\partial_\mu + \mathrm{i}QeA_\mu\right)\psi \tag{6.1.14}$$

So the Lagrangian of QED $\mathcal{L}$ in Eq. (6.1.9) is invariant under local gauge transformation. Such invariance is called local gauge invariance, or local gauge symmetry.

In addition, it can be verified that the self-interaction term for electromagnetic field

$$\mathcal{L}_A = -\frac{1}{4}\left(\partial_\mu A_\upsilon - \partial_\upsilon A_\mu\right)\left(\partial^\mu A^\upsilon - \partial^\upsilon A^\mu\right) \tag{6.1.15}$$

is also invariant under local gauge transformation.

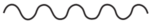
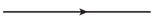
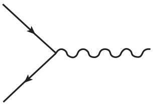
Gauge invariance of QED indicates that not only the overall phase of the fermion field is unmeasurable quantity, but also the whole theory of interaction is invariant because of the existence of electromagnetic field, even if different phases are taken for fermion field at different space–time points.

We can also say it in another words that, if the interaction is required to be local gauge invariant, electromagnetic field must exist. Electromagnetic field is gauge field in this sense. We will see in subsequent chapters that both strong and weak interactions are also interactions described by gauge fields. Equation (6.1.13) is U(1) group transformation, such kind of transformation is also called U(1) gauge transformation.

Some Feynman rules of QED are given in Table 6.1.

Using the Feynman rules, the amplitude of a physical process in electromagnetic interaction can be expanded by the electromagnetic coupling constant, which is called perturbation expansion. Because the electromagnetic coupling constant $\alpha = \frac{e^2}{4\pi} \approx \frac{1}{137}$ is a sufficiently small quantity, it can be expected that this expansion series converges very quickly. In general, the theoretical calculation can be precise enough by only calculating the first few terms in the expansion series. QED can be tested by comparing the theoretical calculation about some processes of electromagnetic interaction and the relevant experimental measurements. At what microscopic scale can the picture of point particle for leptons still reflect the real physical particle can be also tested in this way. Using the scattering processes $\mathrm{e^-e^-} \to \mathrm{e^-e^-}$ and $\mathrm{e^+e^-} \to \mathrm{e^+e^-}$, the comparison of the theoretical calculation with experiment shows that electron can still be viewed as point

Table 6.1. Some Feynman rules of QED.

		Symbol	
Propagator	Photon		$\frac{-\mathrm{i}g_{\mu\nu}}{k^2+\mathrm{i}\varepsilon}$
	Fermion		$\frac{\mathrm{i}}{\not{k}-m+\mathrm{i}\varepsilon}$
Vertex			$\mathrm{i}Qe\gamma_\mu$

particle at the scale $10^{-16} - 10^{-19}$m. At such a microscale, electron still does not show its inner structure.

Several physical processes of electromagnetic interactions shall be discussed in the following sections.

6.2. Møller Scattering

The scattering process $e^-e^- \to e^-e^-$ is called Møller scattering, which is a basic electromagnetic scattering process. QED can be tested to a certain extent by measuring this scattering process in experiment. In 1970s, according to the experimental measurement of this scattering process and the center-of-mass energy used at that experiment 300–500 MeV, it was determined that electron can still be viewed as point particle at the microscale of 10^{-16}m [16].

The leading order Feynman diagram for Møller scattering process is shown in Fig. 6.1, where p_1 and p_2 are the momenta of the initial electrons, and p_3 and p_4 the momenta of the final electrons. Figure 6.1(b) can be obtained from Fig. 6.1(a) by exchanging the position of the two electrons in the final state. According to electron's fermion property, exchanging two electrons will contribute a minus sign in the amplitude, therefore the total amplitude is the subtraction of contributions of Figs. 6.1(a) and 6.1(b). Using the Feynman rule, we can write the scattering amplitude

$$\begin{aligned} A &= \bar{u}_4(-\mathrm{i})e\gamma_\nu u_2 \frac{-\mathrm{i}g^{\mu\nu}}{(p_1-p_3)^2}\bar{u}_3(-\mathrm{i})e\gamma_\mu u_1 \\ &\quad -\bar{u}_4(-\mathrm{i})e\gamma_\mu u_1 \frac{-\mathrm{i}g^{\mu\nu}}{(p_1-p_4)^2}\bar{u}_3(-\mathrm{i})e\gamma_\nu u_2 \\ &= \mathrm{i}e^2\left[\bar{u}_4\gamma^\mu u_2\bar{u}_3\gamma_\mu u_1\frac{1}{(p_1-p_3)^2} - \bar{u}_4\gamma^\mu u_1\bar{u}_3\gamma_\mu u_2\frac{1}{(p_1-p_4)^2}\right] \end{aligned} \tag{6.2.1}$$

The cross section can be calculated with the formulas in Eqs. (3.4.13) and (3.4.14). Let us consider the case of unpolarized electrons, that is, the

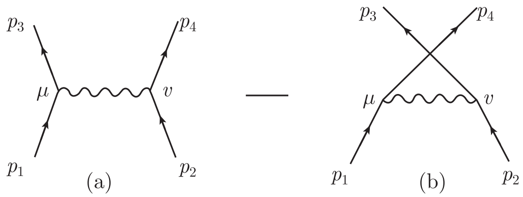


Fig. 6.1. The lowest-order Feynman diagram of Møller scattering.

polarization states of the final electrons are summed and the polarization states of the initial state are averaged.

Then, the differential cross section is

$$\mathrm{d}\sigma = \frac{1}{2}\frac{1}{4}\sum_{\lambda_i}\sum_{\lambda_f}\frac{m_\mathrm{e}^2}{\sqrt{(p_1\cdot p_2)^2-m_\mathrm{e}^4}}|A|^2$$

$$\cdot\frac{\mathrm{d}^3p_3}{(2\pi)^3}\frac{m_\mathrm{e}}{E_3}\frac{\mathrm{d}^3p_4}{(2\pi)^3}\frac{m_\mathrm{e}}{E_4}(2\pi)^4\delta^4(p_1+p_2-p_3-p_4) \qquad (6.2.2)$$

where the first term $\frac{1}{2}$ is the statistic factor for the two identical particles in the final state, and $\frac{1}{4}$ is for averaging the initial polarization states, and $\sum_{\lambda_i}\sum_{\lambda_f}$ is for the sum of the polarization of the initial and final states. Let us deal with the modular square of the amplitude $|A|^2$. Considering the sum of the polarization of the initial and final states, we shall use the following formula:

$$\sum_{\lambda}u^{(\lambda)}(p)\bar{u}^{(\lambda)}(p) = \frac{\not{p}+m}{2m} \qquad (6.2.3)$$

Using this formula, we can obtain

$$\begin{aligned}\sum_{\lambda_i}\sum_{\lambda_f}|A|^2 = e^4\Bigg[&\mathrm{Tr}\left(\frac{\not{p}_2+m_\mathrm{e}}{2m_\mathrm{e}}\gamma^\mu\frac{\not{p}_4+m_\mathrm{e}}{2m_\mathrm{e}}\gamma^\nu\right)\\ &\times\mathrm{Tr}\left(\frac{\not{p}_1+m_\mathrm{e}}{2m_\mathrm{e}}\gamma_\mu\frac{\not{p}_3+m_\mathrm{e}}{2m_\mathrm{e}}\gamma_\nu\right)\frac{1}{(p_1-p_3)^4}\\ &-\mathrm{Tr}\left(\frac{\not{p}_2+m_\mathrm{e}}{2m_\mathrm{e}}\gamma^\mu\frac{\not{p}_4+m_\mathrm{e}}{2m_\mathrm{e}}\gamma^\nu\;\frac{\not{p}_1+m_\mathrm{e}}{2m_\mathrm{e}}\gamma_\mu\frac{\not{p}_3+m_\mathrm{e}}{2m_\mathrm{e}}\gamma_\nu\right.\\ &\left.+\frac{\not{p}_1+m_\mathrm{e}}{2m_\mathrm{e}}\gamma^\mu\frac{\not{p}_4+m_\mathrm{e}}{2m_\mathrm{e}}\gamma^\nu\frac{\not{p}_2+m_\mathrm{e}}{2m_\mathrm{e}}\gamma_\mu\frac{\not{p}_3+m_\mathrm{e}}{2m_\mathrm{e}}\gamma_\nu\right)\\ &\times\frac{1}{(p_1-p_4)^2(p_1-p_3)^2}+\mathrm{Tr}\left(\frac{\not{p}_1+m_\mathrm{e}}{2m_\mathrm{e}}\gamma^\mu\frac{\not{p}_4+m_\mathrm{e}}{2m_\mathrm{e}}\gamma^\nu\right)\\ &\times\mathrm{Tr}\left(\frac{\not{p}_2+m_\mathrm{e}}{2m_\mathrm{e}}\gamma_\mu\frac{\not{p}_3+m_\mathrm{e}}{2m_\mathrm{e}}\gamma_\nu\right)\frac{1}{(p_1-p_4)^4}\Bigg]\end{aligned} \qquad (6.2.4)$$

where the first and the third terms are the contributions of Figs. 6.1(a) and 6.1(b), respectively, and the second term is the interference of this two diagrams. Substituting Eq. (6.2.4) into Eq. (6.2.2), and eliminating m_e in denominator, we get

$$\begin{aligned} \mathrm{d}\sigma = {} & \frac{1}{8}\frac{1}{4\sqrt{(p_1\cdot p_2)^2-m_\mathrm{e}^4}}\left|A'\right|^2\frac{\mathrm{d}^3p_3}{(2\pi)^3}\frac{1}{2E_3}\frac{\mathrm{d}^3p_4}{(2\pi)^3}\frac{1}{2E_4} \\ & \times(2\pi)^4\delta^4(p_1+p_2-p_3-p_4) \end{aligned} \tag{6.2.5}$$

where

$$\begin{aligned} \left|A'\right|^2 = e^4 \Big[& \mathrm{Tr}\,((\not{p}_2+m_\mathrm{e})\gamma^\mu(\not{p}_4+m_\mathrm{e})\gamma^\nu) \\ & \times\mathrm{Tr}\,((\not{p}_1+m_\mathrm{e})\gamma_\mu(\not{p}_3+m_\mathrm{e})\gamma_\nu)\frac{1}{(p_1-p_3)^4} \\ & -\mathrm{Tr}\,((\not{p}_2+m_\mathrm{e})\gamma^\mu(\not{p}_4+m_\mathrm{e})\gamma^\nu\;(\not{p}_1+m_\mathrm{e})\gamma_\mu(\not{p}_3+m_\mathrm{e})\gamma_\nu \\ & +(\not{p}_1+m_\mathrm{e})\gamma^\mu(\not{p}_4+m_\mathrm{e})\gamma^\nu(\not{p}_2+m_\mathrm{e})\gamma_\mu(\not{p}_3+m_\mathrm{e})\gamma_\nu) \\ & \times\frac{1}{(p_1-p_4)^2(p_1-p_3)^2} \\ & +\mathrm{Tr}\,((\not{p}_1+m_\mathrm{e})\gamma^\mu(\not{p}_4+m_\mathrm{e})\gamma^\nu) \\ & \times\mathrm{Tr}\,((\not{p}_2+m_\mathrm{e})\gamma_\mu(\not{p}_3+m_\mathrm{e})\gamma_\nu)\,\frac{1}{(p_1-p_4)^4}\Big] \end{aligned} \tag{6.2.6}$$

To simplify the calculation, considering the mass of electron $m_\mathrm{e}=0.51\mathrm{MeV}$ is a small quantity, we can ignore electron mass at extremely large center-of-mass energy. That is to take $m_\mathrm{e}\to 0$. At such limit, Eqs. (6.2.5) and (6.2.6) become

$$\mathrm{d}\sigma = \frac{1}{32p_1\cdot p_2}\left|A'\right|^2\frac{\mathrm{d}^3p_3}{(2\pi)^3}\frac{1}{2E_3}\frac{\mathrm{d}^3p_4}{(2\pi)^3}\frac{1}{2E_4}(2\pi)^4\,\delta^4\,(p_1+p_2-p_3-p_4) \tag{6.2.7}$$

where $|A'|$ can be simplified as

$$\begin{aligned} \left|A'\right|^2 = e^4 \Bigg[& \mathrm{Tr}\,(\not{p}_2\gamma^\mu\;\not{p}_4\gamma^\nu)\,\mathrm{Tr}\,(\not{p}_1\gamma_\mu\;\not{p}_3\gamma_\nu)\,\frac{1}{(p_1-p_3)^4} \\ & -\mathrm{Tr}\,(\not{p}_2\gamma^\mu\;\not{p}_4\gamma^\nu\;\not{p}_1\gamma_\mu\;\not{p}_3\gamma_\nu+\;\not{p}_1\gamma^\mu\;\not{p}_4\gamma^\nu\;\not{p}_2\gamma_\mu\;\not{p}_3\gamma_\nu) \\ & \times\frac{1}{(p_1-p_4)^2\,(p_1-p_3)^2} \\ & +\mathrm{Tr}\,(\not{p}_1\gamma^\mu\;\not{p}_4\gamma^\nu)\,\mathrm{Tr}\,(\not{p}_2\gamma_\mu\;\not{p}_3\gamma_\nu)\,\frac{1}{(p_1-p_4)^4}\Bigg] \end{aligned} \tag{6.2.8}$$

Using the formulas

$$\gamma^\mu \not{a} \not{b} \not{c} \gamma_\mu = -2 \not{c} \not{b} \not{a}$$

and

$$\gamma^\mu \not{a} \not{b} \gamma_\mu = 4 a \cdot b$$

the second term of Eq. (6.2.8) can be simplified further

$$+16 \mathrm{Tr} \left[\not{p}_2 \not{p}_1 \right] p_3 \cdot p_4 \frac{1}{\left(p_1 - p_4\right)^2 \left(p_1 - p_3\right)^2} \tag{6.2.9}$$

Finally, finish the trace calculation in Eqs. (6.2.8) and (6.2.9), then we can obtain

$$\begin{aligned} |A'|^2 = 32 e^4 & \left[\frac{p_1 \cdot p_2 p_3 \cdot p_4 + p_2 \cdot p_3 p_1 \cdot p_4}{\left(p_1 - p_3\right)^4} + \frac{2 p_1 \cdot p_2 p_3 \cdot p_4}{\left(p_2 - p_4\right)^2 \left(p_1 - p_3\right)^2} \right. \\ & \left. + \frac{p_1 \cdot p_2 p_3 \cdot p_4 + p_1 \cdot p_3 p_2 \cdot p_4}{\left(p_1 - p_4\right)^4} \right] \end{aligned} \tag{6.2.10}$$

Using

$$\frac{1}{2E} = \int \mathrm{d}p^0 \delta\left(p^2 - m^2\right) \theta(p^0)$$

transforming $\frac{\mathrm{d}^3 p_4}{(2\pi)^3} \frac{1}{2E_4}$ in Eq. (6.2.7) to $\mathrm{d}^4 p_4 \delta\left(p_4^2 - m_{\mathrm{e}}^2\right) \theta(E_4) \rightarrow \mathrm{d}^4 p_4 \delta(p_4^2) \theta(E_4)$, and integrating to $\mathrm{d}^4 p_4$, then we can get

$$\mathrm{d}\sigma = \frac{1}{32 p_1 \cdot p_2} |A'|^2 \frac{\mathrm{d}^3 p_3}{(2\pi)^3} \frac{1}{2E_3} \delta \left[\left(p_1 + p_2 - p_3\right)^2 \right] (2\pi) \tag{6.2.11}$$

Next let us calculate the cross section at the center-of-mass frame. Let the energy of the initial electron in the center-of-mass system be E, the momenta are depicted in Fig. 6.2, and θ is the polar angle of $\vec{p}_3$. Using $p_1 + p_2 = (2E, \vec{0})$, we can obtain from (6.2.11)

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{|A'|^2}{32 \times 32 \pi^2 p_1 \cdot p_2} \tag{6.2.12}$$

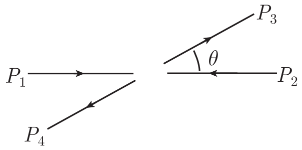


Fig. 6.2.

In the case of ignoring the mass of electron,

$$\begin{aligned} p_1 \cdot p_2 = p_3 \cdot p_4 &= 2E^2 \\ p_1 \cdot p_3 = p_2 \cdot p_4 &= 2E^2 \sin^2 \frac{\theta}{2} \\ p_1 \cdot p_4 = p_2 \cdot p_3 &= 2E^2 \cos^2 \frac{\theta}{2} \end{aligned} \tag{6.2.13}$$

Substitute the result of Eq. (6.2.13) into Eqs. (6.2.10) and (6.2.12), then we get

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{\alpha^2}{16E^2} \left[\frac{1+\cos^4(\theta/2)}{\sin^4(\theta/2)} + \frac{2}{\sin^2(\theta/2)\cos^2(\theta/2)} + \frac{1+\sin^4(\theta/2)}{\cos^4(\theta/2)} \right] \tag{6.2.14}$$

where in Eq. (6.2.14) $\alpha \equiv \frac{e^2}{4\pi}$ is the electromagnetic fine structure constant. Equation (6.2.14) is the angular distribution of the differential cross section of Møller scattering. QED can be tested by measuring the differential cross section and comparing it with Eq. (6.2.14). In 1970s, such experiment was carried out in $\mathrm{e}^-\mathrm{e}^-$ collider [16], with the beam energy of 300 MeV and 550 MeV, respectively. In addition, a deviation term from photon propagator was added during the calculation of the cross section

$$\frac{1}{q^2} \to \frac{1}{q^2} \pm \frac{1}{q^2 - \Lambda^2} \tag{6.2.15}$$

This is equivalent to assuming that in addition to the contribution of the photon propagator in QED, there is an extra contribution of heavy photon-like particle with mass Λ. If Eq. (6.2.15) is transformed into the coordinate space, it is easy to see that this is equivalent to assuming that there is a correction term to Coulomb potential

$$\frac{1}{r} \to \frac{1}{r}\left(1 \pm \mathrm{e}^{-\Lambda r}\right) \tag{6.2.16}$$

Equation (6.2.16) indicates that the larger the value of Λ, the closer the interaction tend to the Coulomb potential, then the result is closer to the prediction of QED. The minimum value of Λ allowed by experiment corresponds to the upper bound of the electron scale $R = 1/\Lambda$. The experimental result was

$$\Lambda > 2.4\mathrm{GeV} \tag{6.2.17}$$

which corresponds to

$$R \leqslant 1 \times 10^{-16}\mathrm{m} \tag{6.2.18}$$

Namely the $\mathrm{e}^-\mathrm{e}^-$ scattering experiment at that time proves that electron can still be viewed as point particle at microscale of 10^{-16} m, and QED is consistent with experiment at the scale of 10^{-16} m.

In 1980s, MARK-J group working on the electron–positron collider PETRA at the electron synchrotron center in Hamburg, Germany, updated the measurement of the radius of electron by the scattering process $\mathrm{e}^+\mathrm{e}^- \to \mathrm{e}^+\mathrm{e}^-$. The radius of electron was updated to be smaller than 10^{-18}m [18]. In 1990s, the electron's radius was measured to be

$$R < 1.4 \times 10^{-19}\mathrm{m} \tag{6.2.19}$$

by the experiment on electron–positron collider (LEP) at CERN [19]. Electron still behaves like a point particle at microscale as small as 10^{-19} m. It does not show its inner structure at such a microscale.

6.3. Bhabha Scattering

The scattering process $\mathrm{e}^+\mathrm{e}^- \to \mathrm{e}^+\mathrm{e}^-$ is called Bhabha scattering. The lowest-order Feynman diagrams for Bhabha scattering are depicted in Fig. 6.3, where (a) is for the one-photon exchange diagram, and (b) the $\mathrm{e}^+\mathrm{e}^-$ annihilation diagram. The contributions of these two diagrams differ by a minus sign. The reason is still because electrons obey Fermi–Dirac statistics. The momentum squared of the photon in Fig. 6.3(a) is $q^2 = (p_1 - p_3)^2 < 0$, such photon is called space-like photon, while for the case of Fig. 6.3(b), the momentum squared of the photon is $q^2 = (p_1 + p_2)^2 > 0$, this kind of photon is called time-like photon. Both space-like and time-like photons are not real photons. The rest mass of a real photon is 0, i.e. $q^2 = 0$. Space-like and time-like photons are virtual photons, which are not real particles because they have no definite mass and definite quantized energy. They are only a state of electromagnetic field. Generally, the electromagnetic field can be in such a state existing for a

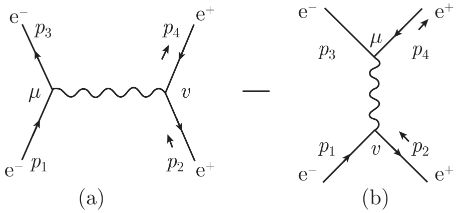


Fig. 6.3. The lowest-order Feynman diagrams for $e^+e^- \to e^+e^-$ scattering process.

very short time. From this point of view, virtual photon is not completely virtual.

According to Feynman rules of QED, the amplitude of Bhabha scattering is

$$\begin{aligned} A &= \bar{u}_3(-\mathrm{i})e\gamma_\mu u_1 \frac{-\mathrm{i}g^{\mu\nu}}{(p_1-p_3)^2}\bar{v}_2(-\mathrm{i})e\gamma_\nu v_4 \\ &\quad -\bar{u}_3(-\mathrm{i})e\gamma_\mu v_4 \frac{-\mathrm{i}g^{\mu\nu}}{(p_1+p_2)^2}\bar{v}_2(-\mathrm{i})e\gamma_\nu u_1 \\ &= \mathrm{i}e^2\left[\bar{u}_3\gamma_\mu u_1\bar{v}_2\gamma^\mu v_4\frac{1}{(p_1-p_3)^2} - \bar{u}_3\gamma_\mu v_4\bar{v}_2\gamma^\mu u_1\frac{1}{(p_1+p_2)^2}\right] \end{aligned} \tag{6.3.1}$$

Similar to Møller scattering of Sec. 6.2, using

$$\sum\nolimits_\lambda u^{(\lambda)}\bar{u}^{(\lambda)} = \frac{\not{p}+m}{2m}, \quad \sum\nolimits_\lambda v^{(\lambda)}\bar{v}^{(\lambda)} = \frac{\not{p}-m}{2m}, \tag{6.3.2}$$

we can finally obtain the differential cross section of $e^+e^- \to e^+e^-$ scattering process

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}(e^+e^- \to e^+e^-) = \frac{\alpha^2}{8E^2}\left[\frac{1+\cos^4(\theta/2)}{\sin^4(\theta/2)} - \frac{2\cos^4(\theta/2)}{\sin^2(\theta/2)} + \frac{1+\cos^2\theta}{2}\right] \tag{6.3.3}$$

where the first term is the contribution of Fig. 6.3(a), the third term the contribution of Fig. 6.3(b), and the second term the interference of Figs. 6.3(a) and 6.3(b). For $e^+e^- \to e^+e^-$ scattering, the contribution of Fig. 6.3(a) dominates, and the secondary major contribution comes from the interference term, because the denominators of the first and second terms are $\sin^4(\theta/2)$ and $\sin^2(\theta/2)$, respectively, which are always smaller than the denominator of the third term for any value of θ. Numerical analysis shows that the contribution of the first term is much greater than that of the

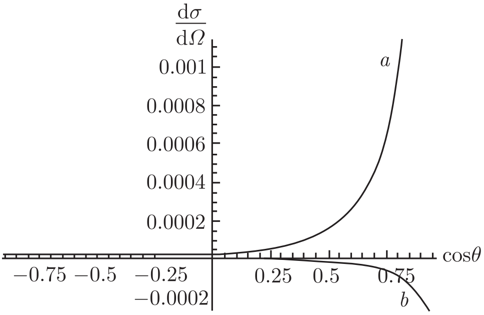


Fig. 6.4. The differential cross-section of $e^+e^- \to e^+e^-$ as a function of $\cos\theta$. The curve (a) is the total contribution of the three terms in Eq. (6.3.3), and (b) the contribution of the second and third terms.

second and third terms for any values of the scattering angle. It is shown in Fig. 6.4 that, the curve (a) is the total contribution of these three terms, and (b) the contributions of the second and third terms. Therefore the main contribution comes from the one-photon exchange diagram of Fig. 6.3(a). Experimentally it is very hard to test the contribution of the annihilation diagram of e^+e^- in QED, i.e. the time-like photon's contribution, through Bhabha scattering process because of the dominance of space-like photon's contribution. Time-like photon's contribution can only be tested through processes of $e^+e^- \to \mu^+\mu^-$ or $e^+e^- \to \bar{q}q$, because there is only contribution of time-like photons in these kinds of processes.

It is also shown in Fig. 6.4 that, the cross section of Bhabha scattering reaches its maximum value when $\cos\theta = 1$, which implies that the scattering probability is large for small scattering angle. Most electrons or positrons scatter in a direction close to the beam axis. Figure 6.3 is the lowest-order Feynman diagram in QED, which only gives the leading-order contribution. With the radiative corrections of QED, corrections of weak and strong interactions, the theoretical calculation are consistent with experiment even better, which is the strong support for QED [17–19].

6.4. The Electromagnetic Form Factors of Nucleons

Experiments of Møller scattering and Bhabha scattering proved that electron still did not show any characteristics of inner structure at the scale of 10^{-19} m. At such scale, electron can still be treated as point particle. Therefore, electron can be used as probe to detect other hadrons' structures. The structure of the nucleon can be determined to a certain extent by hitting the nucleon experimentally with electrons, measuring the scattering cross section of the electron and comparing it with theoretical model. Let us start with the case of electron–nucleon elastic scattering, in which the final particles do not change, they are still electrons and nucleons, but there are momentum transfers, and the momentum and energy of the final electrons and nucleons change.

To distinguish whether nucleon has inner structure or not, let us first calculate the cross section of electron–nucleon elastic scattering assuming that nucleon is a point particle.

6.4.1. *Electron–proton elastic scattering assuming proton being a point-like particle*

The Feynman diagram of electron–proton scattering is depicted in Fig. 6.5, where the black point denotes the interaction vertex of proton and electromagnetic field. If we assume proton is a point particle, then the vertex between proton and electromagnetic field is just the interaction vertex of charged fermion with spin 1/2 and electromagnetic field, which is $\mathrm{i}Qe\gamma_\mu$. For the case of proton, $Q = 1$. Then the amplitude of electron–proton scattering is

$$\begin{aligned} A &= \bar{u}(p')\mathrm{i}e\gamma^\mu u(p)\frac{-\mathrm{i}}{q^2}\bar{u}(k')(-\mathrm{i})e\gamma_\mu u(k) \\ &= -\mathrm{i}e^2\bar{u}(p')\gamma^\mu u(p)\bar{u}(k')\gamma_\mu u(k)\frac{1}{q^2} \end{aligned} \tag{6.4.1}$$

where k and k' are the momenta of electrons in the initial and final states, and p and p' the momenta of protons in the initial and final states, respectively. In the laboratory frame, $p^\mu = (M_\mathrm{p}, \vec{0})$, where M_p is the mass of

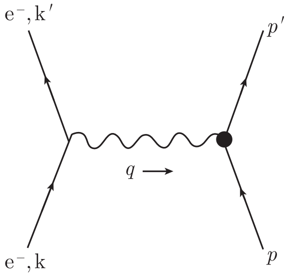


Fig. 6.5. Elastic electron–proton scattering, where the black point denotes the interaction vertex between proton and electromagnetic field.

proton. q is the momentum transfer, $q = k - k'$. In the laboratory frame, we have

$$q^2 = (k - k')^2 = -4EE' \sin^2(\theta/2) \tag{6.4.2}$$

$$(q + p)^2 = p'^2 \Rightarrow q^2 + 2M_\mathrm{p}(E - E') = 0 \tag{6.4.3}$$

where θ is the scattering angle, and E and E' are the energy of the initial and final electrons, respectively. From Eqs. (6.4.2) and (6.4.3), we can obtain

$$E' = \frac{E}{1 + 2\frac{E}{M_\mathrm{p}} \sin^2(\theta/2)} \tag{6.4.4}$$

We have ignored the mass of electron in Eqs. (6.4.2)∼(6.4.4), $m_\mathrm{e} \sim 0$. Using the Trace rule and tricks introduced in Secs. 6.2 and 6.3, we can finally obtain the differential cross section for elastic electron–proton scattering in the case of proton being assumed to be a point particle

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{\alpha^2 \cos^2(\theta/2)}{4E^2 \sin^4(\theta/2)} \frac{1}{1 + \dfrac{2E}{M_\mathrm{p}} \sin^2(\theta/2)} \left[1 + \frac{Q^2}{2M_\mathrm{p}^2} \tan^2(\theta/2)\right] \tag{6.4.5}$$

where $Q^2 \equiv -q^2$.

Proton is a charged particle with spin 1/2 and charge $Q = 1$. When electron is scattered by proton, it is not only affected by the electric field of the proton, but also affected by the proton's magnetic moment. Because charged particles with spin must have intrinsic magnetic moment, so the scattering formula of Eq. (6.4.5) is not only contributed by the electric field of the proton, but also contributed by proton's magnetic moment. In order to distinguish the physical meaning of the first and second terms in the bracket in Eq. (6.4.5), we consider the scattering of electrons on a charged particle with spin of 0 and charge of $Q = 1$. Since the magnetic moment of such charged particle is zero, the scattering on such a particle would be only contribution of the electric field of the charged particle. The interaction vertex of a charged particle of spin 0 and the electromagnetic field is $\mathrm{i}Qe(p+p')_\mu$. Using this vertex and repeating the above calculation, we can finally obtain the differential cross section of the scattering of electron and a charged particle with spin 0 and charge $Q = 1$,

$$\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_\mathrm{Mot} = \frac{\alpha^2 \cos^2(\theta/2)}{4E^2 \sin^4(\theta/2)} \frac{1}{1 + \dfrac{2E}{M_\mathrm{p}} \sin^2(\theta/2)} \tag{6.4.6}$$

This is called Mott (N.F. Mott) scattering formula. Comparing Eqs. (6.4.5) and (6.4.6), we can find that the first term of Eq. (6.4.5) is exactly Mott scattering formula, so the first term of Eq. (6.4.5) comes from the contribution of the electric field of the charged particle to the scattering, then the second term comes from the contribution of magnetic moment of the charged particle.

6.4.2. *Elastic scattering of electron–proton*

The case of a real proton is discussed in this section. Proton is a composite particle consisted of quarks. Its interaction with electromagnetic field are not necessarily described by point-particle interactions. Since the exact form of the interaction vertex of proton and electromagnetic field is not known yet, we will write it as $\mathrm{i}\mathit{\Gamma}_\mu(p', p)$, which is a function of p and p'. With this interaction vertex in hand, the scattering amplitude of electron on a real proton can be written as

$$A = \frac{-\mathrm{i}e^2}{q^2}\bar{u}(k')\gamma^\mu u(k)\bar{u}(p')\mathit{\Gamma}_\mu(p, p')u(p) \tag{6.4.7}$$

If the interaction between electron and proton in Eq. (6.4.7) is interpreted as the interaction between two electromagnetic currents, that is, define two electromagnetic currents

$$j^\mu(k', k) = \bar{u}(k')\gamma^\mu u(k) \tag{6.4.8}$$

and

$$J_\mu(p', p) = \bar{u}(p')\mathit{\Gamma}_\mu(p', p)u(p) \tag{6.4.9}$$

then the scattering amplitude can be written as

$$A = \frac{-\mathrm{i}e^2}{q^2}j^\mu(k', k)J_\mu(p', p) \tag{6.4.10}$$

Since proton is a fermion of spin 1/2, $J_\mu(p', p)$ must be made up of four-momenta and γ matrices. According to relativistic covariance, there are 16 independent 4×4 matrices totally, which are 1, γ^μ, $\sigma^{\mu\nu} \equiv \frac{\mathrm{i}}{2}\left(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu\right)$, $\gamma^\mu\gamma^5$, γ^5. In principle, $J_\mu(p', p)$ can be written as a linear combination of these 16 independent γ matrices, but terms of axial vector and pseudoscalars should not exist in $J_\mu(p', p)$ because of parity conservation of electromagnetic interaction. Both axial vector and pseudoscalars violate parity conservation. So $J_\mu(p',p)$ can be generally written as

$$\begin{aligned} J_\mu(p',p) = \bar{u}(p')\,[&\gamma_\mu\ K_1(q^2) + \mathrm{i}\sigma_{\mu\nu}(p'-p)^\nu K_2(q^2) \\ &+ \mathrm{i}\sigma_{\mu\nu}(p'+p)^\nu K_3(q^2) + (p'-p)_\mu K_4(q^2) \\ &+ (p'+p)_\mu K_5(q^2)]\,u(p) \end{aligned} \tag{6.4.11}$$

where the factors K_i's are only functions of $q^2 \equiv (p'-p)^2$, since they can even be functions of $(p'-p)\cdot(p'+p)$ and $(p'+p)^2$, such two terms can always been written as functions of q^2 and the mass of proton.

The term of $(p'+p)$ in Eq. (6.4.11) can be transformed into combinations of the other terms through the following relations:

$$\bar{u}(p')(p'+p)_\mu u(p) = \bar{u}(p')\,[-\mathrm{i}\sigma_{\mu\nu}(p'-p)^\nu + ({M'}_\mathrm{p} + M_\mathrm{p})\gamma_\mu]\,u(p) \tag{6.4.12}$$

and

$$\bar{u}(p')\mathrm{i}\sigma_{\mu\nu}(p'+p)^\nu u(p) = \bar{u}(p')\,[-(p'-p)_\mu + ({M'}_\mathrm{p} - M_\mathrm{p})\gamma_\mu]\,u(p) \tag{6.4.13}$$

Using Eqs. (6.4.12) and (6.4.13), the independent terms in Eq. (6.4.11) can be greatly decreased. Then Eq. (6.4.11) can be simplified as

$$J_\mu(p',p) = \bar{u}(p')\left[\gamma_\mu F_1(q^2) + \mathrm{i}F_2(q^2)\sigma_{\mu\nu}q^\nu + q_\mu F_3(q^2)\right]u(p) \tag{6.4.14}$$

The conservation of electromagnetic current is not affected by strong interactions, because the $U(1)$ gauge invariance of electromagnetic field is still valid even in the presence of strong interactions. The independent terms in Eq. (6.4.14) will be decreased further because of electromagnetic current conservation,

$$0 = q^\mu J_\mu(p',p) = \bar{u}(p')\left[\not{q}F_1(q^2) + \mathrm{i}F_2(q^2)q^\mu\sigma_{\mu\nu}q^\nu + q^2F_3(q^2)\right]u(p) \tag{6.4.15}$$

Using Dirac equation

$$(\not{p} - m)u(p) = 0 \ \ \text{和} \ \ \bar{u}(p)(\not{p} - m) = 0,$$

the first term of Eq. (6.4.15) becomes $(M'_\mathrm{p} - M_\mathrm{p})F_1(q^2)$, and for elastic scattering $M'_\mathrm{p} = M_\mathrm{p}$, so this term is 0. For the second term, since $\sigma_{\mu\nu}$ is antisymmetric tensor, when it is multiplied with symmetric quantity $q^\mu q^\nu$, the result is 0, therefore Eq. (6.4.15) implies $F_3(q^2) = 0$. Finally, only two independent factors in Eq. (6.4.15) survive,

$$J_\mu(p', p) = \bar{u}(p') \left[\gamma_\mu F_1(q^2) + \mathrm{i} F_2(q^2)\sigma_{\mu\nu} q^\nu\right] u(p) \tag{6.4.16}$$

where $F_1(q^2)$ and $F_2(q^2)$ are the form factors for nucleons.

Substitute Eqs. (6.4.8) and (6.4.16) into (6.4.10), and after a few steps of calculation, we can get the cross section of electron–proton elastic scattering,

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_{\mathrm{Mot}} \left[\left(F_1^2 + Q^2 F_2^2\right) + \frac{Q^2}{2M_{\mathrm{p}}^2}\left(F_1 + 2M_{\mathrm{p}} F_2\right)^2 \tan^2\frac{\theta}{2}\right] \tag{6.4.17}$$

where $Q^2 \equiv -q^2$, and M_{p} is the mass of proton. This formula is called Rosenbluth (M.N. Rosenbluth) formula [26].

In the case of small q^2, i.e., the small momentum transfer, the scattering of electron is not sensitive to the inner structure of the proton, so the scattering of electron on the proton is very similar to point-particle scattering. Therefore, in the limit $q^2 \to 0$, Eq. (6.4.17) should return back to the scattering formula of Eq. (6.4.5), which is for the scattering of electron on a point-like proton. Comparing these two formulas, we can obtain the behavior of the form factor $F_1(q^2)$ in the limit $q^2 \to 0$

$$F_1(0) = 1 \tag{6.4.18}$$

The second term in Eq. (6.4.5) is the contribution of the proton's magnetic moment. Relevantly, the second term in Eq. (6.4.17) $\left(F_1 + 2M_{\mathrm{p}} F_2\right)^2$ should also be the contribution of the proton's magnetic moment. Since $F_1(0) = 1$, which is already equal to the contribution of the magnetic moment of a particle of spin 1/2, so we can interpret the term of $2M_{\mathrm{p}}F_2$ as contribution of proton's anomalous magnetic moment. It is obtained in experiment [20],

$$F_2(0) = \frac{1.79}{2M_{\mathrm{p}}} \tag{6.4.19}$$

Equation (6.4.17) shows that, as independent factors, none of $F_1(q^2)$ and $F_2(q^2)$ is solely relevant to magnetic moment contribution or electric contribution to the scattering. However, we can define a new pair of form factors which can be relevant to magnetic moment and electric contribution to the scattering, respectively [27, 28],

$$\begin{aligned} G_{\mathrm{M}} &= F_1(q^2) + 2M_{\mathrm{p}}F_2(q^2) \\ G_{\mathrm{E}} &= F_1(q^2) - \frac{Q^2}{2M_{\mathrm{p}}}F_2(q^2) \end{aligned} \tag{6.4.20}$$

where G_{M} is magnetic form factor, and G_{E} electric form factor.

With the new form factors, Rosenbluth formula can be reexpressed as

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_{\mathrm{Mot}} \left[\frac{G_{\mathrm{E}}^2 + \tau G_{\mathrm{M}}^2}{1+\tau} + 2\tau G_{\mathrm{M}}^2 \tan^2\frac{\theta}{2}\right] \tag{6.4.21}$$

where $\tau \equiv Q^2/4M_{\mathrm{p}}^2$.

The physical explanation of the form factors is discussed below.

Let us first consider the problem of the electron being scattered by a field potential of a certain charge distribution. In the non-relativistic case, the differential cross section of an electron scattered by a potential $U(r)$ generated by a certain charge distribution can be obtained by Born approximation,

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{m_{\mathrm{e}}^2}{4\pi^2}\left|-\mathrm{e}\int U(r)\mathrm{e}^{\mathrm{i}(\vec{k}-\vec{k}')\cdot\vec{r}}\mathrm{d}^3r\right|^2 \tag{6.4.22}$$

The potential satisfies Poisson equation,

$$\nabla^2 U(r) = -4\pi\rho(r)$$

where $\rho(r)$ is the charge density of the potential.

Define $\vec{q} = \vec{k} - \vec{k}'$, then Eq. (6.4.22) can be written as

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{m_{\mathrm{e}}^2}{4\pi^2}\left|-\mathrm{e}\int U(r)\mathrm{e}^{\mathrm{i}\vec{q}\cdot\vec{r}}\mathrm{d}^3r\right|^2 \tag{6.4.23}$$

where $\int U(r)\mathrm{e}^{\mathrm{i}\vec{q}\cdot\vec{r}}\mathrm{d}^3r$ can be further expressed as

$$\begin{aligned} \int U(r)\mathrm{e}^{\mathrm{i}\vec{q}\cdot\vec{r}}\mathrm{d}^3r &= -\frac{1}{|\vec{q}|^2}\int U(r)\nabla^2\mathrm{e}^{\mathrm{i}\vec{q}\cdot\vec{r}}\mathrm{d}^3r \\ &= -\frac{1}{|\vec{q}|^2}\int [\nabla^2 U(r)]\mathrm{e}^{\mathrm{i}\vec{q}\cdot\vec{r}}\mathrm{d}^3r \\ &= \frac{4\pi}{|\vec{q}|^2}\int \rho(r)\mathrm{e}^{\mathrm{i}\vec{q}\cdot\vec{r}}\mathrm{d}^3r \end{aligned} \tag{6.4.24}$$

Substitute the above result into Eq. (6.4.23), and use $|\vec{q}| = 2\left|\vec{k}\right|\sin\frac{\theta}{2}$, then we can get

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{m_{\mathrm{e}}^2 e^4}{4|\vec{k}|^4 \sin^4(\theta/2)} \left|\frac{1}{e}\int \rho(r)\mathrm{e}^{\mathrm{i}\vec{q}\cdot\vec{r}}\mathrm{d}^3 r\right|^2 \tag{6.4.25}$$

where

$$\frac{1}{e}\int \rho(r)\mathrm{e}^{\mathrm{i}\vec{q}\cdot\vec{r}}\mathrm{d}^3 r \tag{6.4.26}$$

can be defined as the form factors for the potential, that is

$$F^E(q^2) \equiv \frac{1}{e}\int \rho(r)\mathrm{e}^{\mathrm{i}\vec{q}\cdot\vec{r}}\mathrm{d}^3 r \tag{6.4.27}$$

The physical interpretation of the form factor is obtained, that is, the form factor is the Fourier transformation of the charge density distribution.

For spherically symmetric charge distribution, Eq. (6.4.27) can be expanded as

$$F^{\mathrm{E}}(q^2) = 1 - \frac{|\vec{q}|^2}{6}\left\langle r^2\right\rangle + \cdots \tag{6.4.28}$$

where $\left\langle r^2\right\rangle \equiv \frac{1}{e}\int \rho(r) r^2 \mathrm{d}^3 r$, and the charge that generates the potential is assumed to be

$$\frac{1}{e}\int \rho(r)\mathrm{d}^3 r = 1 \tag{6.4.29}$$

So the average of the radius squared of the charge distribution is

$$\left\langle r^2\right\rangle = -6\frac{\mathrm{d}F^{\mathrm{E}}(q^2)}{\mathrm{d}\,|\vec{q}|^2}\Big|_{|\vec{q}|^2\to 0} \tag{6.4.30}$$

Equation (6.4.30) is obtained under the non-relativistic limit. In the non-relativistic limit, it can be considered that the scattering center is in a static state, the energy of the electron before and after scattering is unchanged, and only the momentum changes, that is, the four-momentum transfer is $q^\mu = (0, \vec{q})$. With this condition, Eq. (6.4.30) can be written in a relativistic covariant form

$$\left\langle r^2\right\rangle = -6\frac{\mathrm{d}F^{\mathrm{E}}(q^2)}{\mathrm{d}(-q^2)}\Big|_{q^2\to 0} \tag{6.4.31}$$

This result corresponds to the charge distribution of the potential field in coordinate space in the non-relativistic limit.

From the above discussion, we can know that the electric form factor of nucleon corresponds to the charge distribution of nucleon, while the magnetic form factor corresponds to the magnetic moment distribution of the nucleon. The radius squared mean value of the charge distribution and the magnetic moment distribution of the proton given in the experiment are

$$\begin{aligned}\left\langle r^2\right\rangle_{\mathrm{E}} &= -6\frac{\mathrm{d}G_{\mathrm{E}}(Q^2)}{\mathrm{d}Q^2}|_{Q^2\to 0} \approx 0.64\times 10^{-30}\,\mathrm{m}^2 \\ \left\langle r^2\right\rangle_{\mathrm{M}} &= -6\frac{\mathrm{d}G_{\mathrm{M}}(Q^2)}{\mathrm{d}Q^2}|_{Q^2\to 0} \approx 0.64\times 10^{-30}\,\mathrm{m}^2\end{aligned} \tag{6.4.32}$$

where $Q^2 = -q^2$.

The experimental measurement on the nucleon electromagnetic form factor indicates that nucleon is not a point particle, and there are certain distributions for its charge and magnetic moment. Some forms of charge distribution and the corresponding form factors and $\left\langle r^2\right\rangle$ are listed in Table 6.2.

There is no fully solvable theory for the structure of hadrons at present, so the form factor of nucleon cannot be calculated exactly in theory up to now. Now the variation of the from factor with Q^2 is fitted according to the experimental data. The electromagnetic form factor of proton in elastic scattering is determined by the elastic scattering of electron on the liquid hydrogen target, while the electromagnetic from factor of neutron is difficult to be determined directly. Experimentally, electron–deuteron and electron–proton elastic scattering are measured separately, and then subtract between them to get the cross section of electron–neutron elastic

Table 6.2. Some charge distribution and corresponding form factors and mean square radii $\left\langle r^2\right\rangle$.

Charge distribution $\rho(r)$	Form factor	Mean square radius $\left\langle r^2\right\rangle$
Point-charge $\delta(\vec{r})$	1	0
Yukawa distribution $\frac{\alpha^2}{4\pi}\frac{\mathrm{e}^{-\alpha r}}{r}$	single pole $\frac{1}{1+\lvert\vec{q}\rvert^2/\alpha^2}$	$6/\alpha^2$
Exponential distribution $\frac{m^3}{8\pi}\mathrm{e}^{-mr}$	dipole distribution $\frac{1}{\left(1+\lvert\vec{q}\rvert^2/m^2\right)^2}$	$12/m^2$

scattering,

$$\frac{\mathrm{d}\sigma(\mathrm{en})}{\mathrm{d}\Omega} = \frac{\mathrm{d}\sigma(\mathrm{ed})}{\mathrm{d}\Omega} - \frac{\mathrm{d}\sigma(\mathrm{ep})}{\mathrm{d}\Omega} \tag{6.4.33}$$

But the result measured in this way is not so precise, because some terms about the structure of deuteron need to be added into Eq. (6.4.33). Even in this way the result is still not precise enough. But experimental data in hand have shown that neutron is not a point particle. The total charge of neutron is 0, and no evidence of positive and negative charge distribution have been indicted in experiment. The electric form factor of neutron is zero, but its magnetic form factor is not zero, which implies that there is a certain magnetic moment distribution in neutron. Since 1961, many experiments have shown the electromagnetic form factors of proton and neutron satisfy the following relation [29]:

$$\begin{aligned} &G_{\mathrm{E}}^{\mathrm{p}}(Q^2) = \frac{G_{\mathrm{M}}^{\mathrm{p}}(Q^2)}{\mu_{\mathrm{p}}} = \frac{G_{\mathrm{M}}^{\mathrm{n}}(Q^2)}{\mu_{\mathrm{n}}} \equiv G(Q^2), \\ &G_{\mathrm{E}}^{\mathrm{n}}(Q^2) = 0 \end{aligned} \tag{6.4.34}$$

where $\mu_{\mathrm{p}} \equiv G_{\mathrm{M}}^{\mathrm{p}}(0)$, $\mu_{\mathrm{n}} \equiv G_{\mathrm{M}}^{\mathrm{n}}(0)$. Experimentally, $G(Q^2)$ can be well fitted with the dipole form for the form factors

$$G(Q^2) = \frac{1}{\left(1 + \dfrac{Q^2}{0.71}\right)^2} \tag{6.4.35}$$

where Q^2 is in the unit of GeV^2. Dipole form of the form factor implies that the charge and magnetic moment distribution of the proton is exponential form, which can be shown by performing Fourier transformation for Eq. (6.4.35),

$$\rho(r) = \rho_0 \mathrm{e}^{-\sqrt{0.71}r} \tag{6.4.36}$$

The mean square radius can be calculated using Eq. (6.4.36)

$$\sqrt{\langle r^2 \rangle} = \sqrt{\frac{12}{0.71}} \approx 0.8 \text{ fm} \tag{6.4.37}$$

This can be viewed as the scale of the nucleon measured in experiment, that is, the root-mean-square (RMS) radius of the nucleon is about 0.8 fm.

6.5. Inelastic Scattering of Electron–Proton

Hitting nucleons with electrons, when the energy of electron is low, it is not possible to convert the nucleon into higher excited state or break it up. What happens in this case is electron–nucleon elastic scattering. The experiment of electron–nucleon elastic scattering shows that nucleon is a particle with a certain charge and magnetic moment distribution, and its distribution radius is about 0.8 fm. When the transferred energy and momentum of electron in the scattering increase, the nucleon may transform into a series of excited states. At this time, a series of peak structures may be found in experiment by measuring the invariant mass spectrum of the final hadrons. Such reactions are inelastic scattering processes, such as the process

$$\begin{tikzcd}[column sep=small] \mathrm{e}+\mathrm{p}\to\mathrm{e}+ & \Delta^+ \arrow[dr, to path={|- (\tikztotarget)}] & \\ & & \pi^0+\mathrm{p} \text{ or } \pi^+ + \mathrm{n} \end{tikzcd}$$

The invariant mass of the hadrons in this process is $M_{\mathrm{X}} = 1232\,\mathrm{MeV}$. If the transferred energy and momentum increase further, the nucleon may break up into large number of hadrons. At such time, there are no peaks in the invariant mass spectrum. Such scattering is called inelastic electron–nucleon scattering. Since the momentum transfer squared $\left|q^2\right|$ is large at this time, according to the uncertainty principle of quantum mechanics, the resolution of electron is extremely high as a probe. The inner structure of nucleon can be detected with this probe. In the 1960s, it was through deep inelastic scattering of lepton–nucleon, our understanding of nucleon structure was greatly improved.

6.5.1. *Structure function of inelastic scattering*

The inelastic scattering process of electron and nucleon is shown in Fig. 6.6. After hit by electron, the nucleon is turned into large number of hadrons. If not all the final hadrons are detected in experiment, only electron in the final state detected, such process is called inclusive process. On the contrary, if each particle in the final state is identified, the process is called exclusive process. Let us denote the four-momentum of the initial electron as k, the energy as E, and the four-momentum of the final electron as k', the energy E'. The four-momentum of the initial nucleon is denoted to be p, the final hadrons to be X totally, and the four-momentum of final hadrons is $p\prime$. q is the momentum transfer $q = k - k' = p' - p$. The invariant mass of the final hadrons is $M_{\mathrm{X}}^2 = p'^2$. The inelastic scattering process of

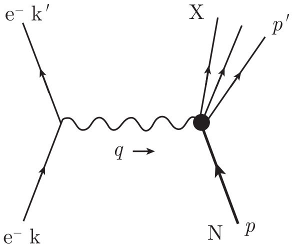


Fig. 6.6. The inelastic scattering process of electron–nucleon.

electron–nucleon is

$$\mathrm{e} + \mathrm{N} \rightarrow \mathrm{e} + \mathrm{X} \tag{6.5.1}$$

where N is the nucleon, which can either be proton p or neutron n.

The left side of the diagram in Fig. 6.6 is the interaction vertex of electron and electromagnetic field. Since electron can still be treated as point particle at energy scale that the experiment can reach, the interaction between electron and electromagnetic field is just described by the vertex of electromagnetic interaction in QED. The electromagnetic coupling constant e is a small quantity in QED, so the first-order approximation of QED is a good approximation. The right side is the interaction vertex of the nucleon, hadron and electromagnetic field. When q^2 is in the range of deep inelastic scattering, the nucleon cannot be treated as point particle. The interaction in this case must involve the details of the nucleon structure. Although the details of the inner structure of the nucleon are not fully understood, the eN inelastic scattering amplitude can be written in the first-order approximation in QED as

$$A_{\mathrm{NE}} = e^2 j_\mu \frac{1}{q^2} J^\mu \tag{6.5.2}$$

where j_μ is the electromagnetic current of electron,

$$j_\mu = \bar{u}(k')\gamma_\mu u(k) \tag{6.5.3}$$

here $1/q^2$ is the photon's propagator, J^μ the electromagnetic current for hadrons. For the process in Eq. (6.5.1), if electron is unpolarized, the differential cross section is proportional to $|A_{\mathrm{NE}}|^2$, i.e.,

$$\mathrm{d}\sigma \propto L_{\mu\nu} W^{\mu\nu} \tag{6.5.4}$$

where $L_{\mu\nu}$ is the lepton current tensor,

$$L_{\mu\nu} \equiv \frac{1}{2}\sum_{\lambda_1}\sum_{\lambda_2} j_\mu^+ j_\nu = 2\left(k_\mu k'_\nu + k_\nu k'_\mu + \frac{q^2}{2}g_{\mu\nu}\right) \tag{6.5.5}$$

with $\frac{1}{2}\sum_{\lambda_1}$ being the average of the initial electron's polarization states, and $\sum_{\lambda_2}$ the sum of the polarization states of the final electron.

$W^{\mu\nu}$ is the unpolarized hadronic current tensor. Its definition is

$$W^{\mu\nu} \equiv \frac{1}{2}\int \mathrm{d}^3p' \sum_{\lambda_3}\sum_{\lambda_4} J^{+\mu} J^\nu \delta^4\left(p+q-p'\right) \tag{6.5.6}$$

where $\frac{1}{2}\sum_{\lambda_3}$ is the average for the polarization states of the initial nucleon, and $\sum_{\lambda_4}$ the sum for the polarization states of the final hadrons. The δ function needed in the calculation of the cross section is also included in the definition of $W^{\mu\nu}$, therefore δ function need not to be duplicately considered in the calculation of the cross section later.

To ensure the Lorentz covariance of the scattering cross section, the hadronic current tensor $W^{\mu\nu}$ must be a combination of the independent momenta existing in the scattering process, and the possible combinations are

$$g^{\mu\nu}, \quad p^\mu p^\nu, \quad p^\mu q^\nu, \quad q^\mu p^\nu, \quad q^\mu q^\nu, \quad \varepsilon^{\mu\nu\rho\sigma}p_\rho q_\sigma \tag{6.5.7}$$

It can be known from Eq. (6.5.5) that $L_{\mu\nu}$ is symmetric for its indices μ and ν. So the antisymmetric terms for indices μ and ν in $W^{\mu\nu}$ do not contribute to the cross section $\mathrm{d}\sigma \propto L_{\mu\nu}W^{\mu\nu}$. Only symmetric terms need to be considered, then

$$W^{\mu\nu} = -W_1 g^{\mu\nu} + W_2\frac{p^\mu p^\nu}{M_\mathrm{N}^2} + W_3\frac{p^\mu q^\nu + q^\mu p^\nu}{M_\mathrm{N}^2} + W_4\frac{q^\mu q^\nu}{M_\mathrm{N}^2} \tag{6.5.8}$$

where W_1, W_2, W_3 and W_4 are scalar functions of p and q. All the possible Lorentz scalars composed of p and q are p^2, q^2 and $p\cdot q$. But $p^2 = M_\mathrm{N}^2$ is not a variable. So W_i's can only be functions of q^2 and $p\cdot q$. Generally it is not $p\cdot q$, but

$$\nu \equiv \frac{p\cdot q}{M_\mathrm{N}} \tag{6.5.9}$$

taken as one of the variables. So W_i's are functions of q^2 and ν.

Using electromagnetic current conservation $q_\mu W^{\mu\nu} = 0$, we can obtain

$$\left[-W_1 + \frac{W_3}{M_{\mathrm{N}}^2} p \cdot q + \frac{W_4}{M_{\mathrm{N}}^2} q^2\right] q^\nu + \left[\frac{W_2}{M_{\mathrm{N}}^2} p \cdot q + \frac{W_3}{M_{\mathrm{N}}^2} q^2\right] p^\nu = 0 \qquad (6.5.10)$$

Because p and q are independent, we have

$$W_3 = -\frac{p \cdot q}{q^2} W_2 \qquad (6.5.11)$$

$$W_4 = -\frac{p \cdot q}{q^2} W_3 + W_1 \frac{M_{\mathrm{N}}^2}{q^2} \qquad (6.5.12)$$

Substituting Eqs. (6.5.11) and (6.5.12) into (6.5.8), we can finally obtain

$$W^{\mu\nu} = \left(\frac{q^\mu q^\nu}{q^2} - g^{\mu\nu}\right) W_1(\nu, q^2) + \left(p^\mu - \frac{p \cdot q}{q^2} q^\mu\right) \left(p^\nu - \frac{p \cdot q}{q^2} q^\nu\right) \frac{W_2(\nu, q^2)}{M_{\mathrm{N}}^2} \qquad (6.5.13)$$

where W_1 and W_2 have been explicitly expressed as functions of ν and q^2. Using Eqs. (6.5.5) and (6.5.13), we can calculate the differential cross section of eN inelastic scattering. After calculation, we can obtain

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega \mathrm{d}E'} = \frac{\alpha^2}{4E^2 \sin^4(\theta/2)} \left[W_2(\nu, q^2) \cos^2 \frac{\theta}{2} + 2W_1(\nu, q^2) \sin^2 \frac{\theta}{2}\right] \qquad (6.5.14)$$

We can also express the differential cross section as function of q^2 and the invariant mass squared of the hadrons M_{X}^2

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}q^2 \mathrm{d}M_{\mathrm{X}}^2} = \frac{2\pi\alpha^2}{M_{\mathrm{N}} q^4} \frac{E'}{E} \left[W_2(\nu, q^2) \cos^2 \frac{\theta}{2} + 2W_1(\nu, q^2) \sin^2 \frac{\theta}{2}\right] \qquad (6.5.15)$$

The functions W_1 (ν, q^2) and W_2 (ν, q^2) in Eqs. (6.5.14) and (6.5.15) are called the structure functions of the nucleons. They reflect the information about the nucleon structure. In principle, as long as the structure of nucleon is known, the structure functions can be calculated in theory. In Sec. 6.6, we shall see that many properties of the structure functions can indeed be derived from the parton model for the nucleons, which are well consistent with experiment.

6.5.2. *The structure function applied to elastic scattering*

Now let us discuss the relation between the nucleon structure function and the nucleon form factors in elastic scattering. Repeat the calculation of the differential cross section of electron–nucleon elastic scattering using Eqs. (6.4.8), (6.4.16) and (6.4.10). The difference from Eq. (6.4.21) is that we do not integrate to $\mathrm{d}E'$ this time. That result is

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega\mathrm{d}E'} = \frac{\alpha^2}{4E^2\sin^4(\theta/2)}\delta\left(\nu - \frac{Q^2}{2M_\mathrm{N}}\right)\left[\frac{G_\mathrm{E}^2 + (Q^2/4M_\mathrm{N}^2)G_\mathrm{M}^2}{1 + (Q^2/4M_\mathrm{N}^2)}\cos^2\frac{\theta}{2} + \frac{Q^2}{2M_\mathrm{N}^2}G_\mathrm{M}^2\sin^2\frac{\theta}{2}\right] \tag{6.5.16}$$

where $Q^2 \equiv -q^2$. Comparing Eqs. (6.5.14) and (6.5.16), we can obtain the structure functions of nucleon in the case of elastic scattering

$$W_1^E = \frac{Q^2}{4M_\mathrm{N}^2}G_\mathrm{M}^2\delta\left(\nu - \frac{Q^2}{2M_\mathrm{N}}\right) \tag{6.5.17}$$

$$W_2^E = \frac{G_\mathrm{E}^2 + (Q^2/4M_\mathrm{N}^2)G_\mathrm{M}^2}{1 + (Q^2/4M_\mathrm{N}^2)}\delta\left(\nu - \frac{Q^2}{2M_\mathrm{N}}\right) \tag{6.5.18}$$

Unlike the case of inelastic scattering, the two variables ν and Q^2 of the structure functions are not independent free variables, they are correlated by the δ function. At this time, the structure function is no longer the function of ν and Q^2, but only function of a single variable ν or Q^2. Write Eqs. (6.5.17) and (6.5.18) as follows:

$$W_1^E = \frac{G_\mathrm{M}^2}{4M_\mathrm{N}^2}\frac{Q^2}{\nu}\delta\left(1 - \frac{Q^2}{2M_\mathrm{N}\nu}\right) \tag{6.5.19}$$

$$W_2^E = \frac{G_\mathrm{E}^2 + (Q^2/4M_\mathrm{N}^2)G_\mathrm{M}^2}{1 + (Q^2/4M_\mathrm{N}^2)}\frac{1}{\nu}\delta\left(1 - \frac{Q^2}{2M_\mathrm{N}\nu}\right) \tag{6.5.20}$$

The above two equations show that, as long as G_E and G_M do not apparently depend on Q^2 when Q^2 and ν are large enough, W_1^E and νW_2^E shall only be functions of Q^2/ν.

After analyzing the experimental data of deep inelastic scattering, Bjorken obtained by using the method of current algebra: for deep inelastic scattering, when both Q^2 and ν are very large, the structure functions $W_1(\nu, Q^2)$ and $\nu W_2(\nu, Q^2)$ will no longer be functions of two free variables ν and Q^2, but only functions of a single scaling variable $x \equiv \frac{Q^2}{2M_\mathrm{N}\nu}$. This is named as scaling phenomenon [24].

Experimental data showed that, for fixed value of $x = \frac{Q^2}{2M_{\mathrm{N}}\nu}$, the structure functions $W_1(\nu, Q^2)$ and $\nu W_2(\nu, Q^2)$ indeed do not simply depend on Q^2, but only depend on x.

According to the discussion on electron–proton elastic scattering in Sec. 6.4.1, and Eqs. (6.5.19) and (6.5.20), we can know that, if nucleon were a point particle of spin 1/2, then νW_2^E shall not depend on Q^2. Therefore, experimental data indicate that electron may be scattered by point-like particles in nucleons in deep inelastic scattering process. This issue will be discussed further in Sec. 6.6.

6.6. The Parton Model for Nucleons

Experiment of electron–nucleon deep inelastic scattering showed the scaling phenomenon for the structure functions, which is the important evidence that there is inner structure in nucleon. These experimental data aroused heated investigation on the nucleon structure at that time, among which the parton model proposed by Feynman is the most outstanding one [25].

It is assumed in the parton model that nucleon is composed of a large number of partons, which are point-like particles with spin of 1/2. Assume that, in the coordinate system where the nucleon's momentum p tends to be infinity, the longitudinal momentum of each parton is x_i times of the nucleon momentum,

$$\left(\vec{p}_{\parallel}\right)_i = x_i \vec{p} \tag{6.6.1}$$

where $0 \leqslant x_i \leqslant 1$, and $\sum_i x_i = 1$. Denote the transverse momentum of the ith parton by $(p_{\perp})_i$, and its mass by m_i. The transverse momenta satisfy the condition

$$\sum_i (p_{\perp})_i = 0 \tag{6.6.2}$$

The energy E_i of each parton may fluctuate in a certain extent, so the total energy of all the partons $E_{\mathrm{V}} = \sum_i E_i$ may not be necessarily equal to the energy of the nucleon E_{N}. According to the uncertainty principle of quantum mechanics, the lifetime of such kind of state τ_{life} is inversely proportional to the energy fluctuation, i.e.

$$\tau_{\mathrm{life}} = \frac{1}{|E_{\mathrm{V}} - E_{\mathrm{N}}|} \tag{6.6.3}$$

In the coordinate system of infinity momentum

$$\begin{aligned}E_{\mathrm{V}}-E_{\mathrm{N}} &= \sum_i \sqrt{\left(\left|\vec{p}_{||}\right|\right)_i^2+(p_\perp)_i^2+m_i^2}-\sqrt{|\vec{p}|^2+M_{\mathrm{N}}^2}\\ &= \sum_i \sqrt{x_i^2\,|\vec{p}|^2+(p_\perp)_i^2+m_i^2}-\sqrt{|\vec{p}|^2+M_{\mathrm{N}}^2}\\ &\approx \sum_i x_i\,|\vec{p}|+\sum_i \frac{(p_\perp)_i^2+m_i^2}{2x_i\,|\vec{p}|}-|\vec{p}|-\frac{M_{\mathrm{N}}^2}{2\,|\vec{p}|}\\ &= \frac{1}{2\,|\vec{p}|}\left(\sum_i \frac{(p_\perp)_i^2+m_i^2}{x_i}-M_{\mathrm{N}}^2\right)\end{aligned} \tag{6.6.4}$$

and the interacting time τ_{int} for each scattering is inversely proportional to q^0,

$$q^0=(1/\gamma)\,q^0_{\mathrm{lab}}=\frac{M_{\mathrm{N}}}{E_{\mathrm{N}}}\nu \tag{6.6.5}$$

where $\gamma=\frac{1}{\sqrt{1-\nu^2/c^2}}$ is the Lorentz transformation factor in theory of relativity. Therefore the interacting time is

$$\tau_{\mathrm{int}}=\frac{1}{q^0}=\frac{E_{\mathrm{N}}}{M_{\mathrm{N}}\nu}=\frac{\sqrt{|\vec{p}|^2+M_{\mathrm{N}}^2}}{M_{\mathrm{N}}\nu}\approx\frac{|\vec{p}|}{M_{\mathrm{N}}\nu} \tag{6.6.6}$$

Compare Eqs. (6.6.6) and (6.6.3), we obtain

$$\frac{\tau_{\mathrm{int}}}{\tau_{\mathrm{life}}}=\frac{1}{2M_{\mathrm{N}}\nu}\left|\sum_i \frac{(p_\perp)_i^2+m_i^2}{x_i}-M_{\mathrm{N}}^2\right| \tag{6.6.7}$$

Equation (6.6.7) indicates that, as long as $\nu\to\infty$, the interacting time shall be seriously smaller than the existing time of the intermediate state. Therefore, in the case that ν is large, the partons in the nucleon can be regarded as quasi-free particles in deep inelastic scattering process. Equation (6.6.7) also shows that, to satisfy $\tau_{\mathrm{int}}\ll\tau_{\mathrm{life}}$, x_i should not be too small, so the contribution from the region of small x_i in the deep inelastic scattering should be treated with special care. The following discussion is restricted to regions where x_i is not too small.

Since in the deep inelastic scattering process, nucleon can be regarded as being composed of quasi-free partons, so the scattering between electron and nucleon can be equivalent to the superposition of electron–parton scattering. Assume that partons are point-like particles with spin of 1/2, the interaction vertex of which with electromagnetic field is $\mathrm{i}Q_i e\gamma_\mu$, where Q_i is the charge of the ith parton. Similar to the calculation procedure of the scattering cross section of electron and point particle of spin 1/2 discussed in Sec. 6.4.1, we can obtain the differential cross section of electron and the ith parton scattering,

$$\frac{\mathrm{d}^2\sigma_i}{\mathrm{d}\Omega \mathrm{d}E'} = \frac{\alpha^2 Q_i^2}{4E^2\sin^4(\theta/2)}\frac{x_i}{\nu}\delta\left(x_i - \frac{Q^2}{2M_{\mathrm{N}}\nu}\right)\left[\cos^2\frac{\theta}{2} + \frac{Q^2}{2M_{\mathrm{N}}^2}\frac{1}{x_i^2}\sin^2\frac{\theta}{2}\right] \tag{6.6.8}$$

In calculation of Eq. (6.6.8), we have used

$$p_i = x_i p, \quad m_i \approx x_i M_{\mathrm{N}} \tag{6.6.9}$$

in the coordinate system of infinity momentum.

Let $f_i(x_i)$ be the probability density of the ith parton's momentum fraction located in the range $x_i \to x_i + \mathrm{d}x_i$, then the cross section of electron–nucleon deep inelastic scattering is

$$\begin{aligned}\frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega \mathrm{d}E'} &= \sum_i \int \mathrm{d}x_i f_i(x_i)\frac{\mathrm{d}^2\sigma_i}{\mathrm{d}\Omega \mathrm{d}E'} \\ &= \frac{\alpha^2}{4E^2\sin^4(\theta/2)}\frac{1}{\nu}\sum_i x f_i(x) Q_i^2\left[\cos^2\frac{\theta}{2} + \frac{Q^2}{2M_{\mathrm{N}}^2}\frac{1}{x^2}\sin^2\frac{\theta}{2}\right]\end{aligned} \tag{6.6.10}$$

where $x = \frac{Q^2}{2M_{\mathrm{N}}\nu}$ is the scaling variable of the electron–nucleon deep inelastic scattering. Equation (6.6.10) also gives another interpretation to the scaling variable x, that is, in the coordinate system of infinity momentum, the scaling variable $x \equiv \frac{Q^2}{2p\cdot q} = \frac{Q^2}{2M_{\mathrm{N}}\nu}$ is equal to the fraction of the momentum carried by the scattering parton relative to the total momentum of nucleon.

Comparing Eqs. (6.6.10) and (6.5.14), we can obtain the following relations:

$$\sum_i \frac{1}{\nu} x f_i(x) Q_i^2 = W_2 \tag{6.6.10a}$$

$$\sum_i \frac{1}{\nu} x f_i(x) Q_i^2 \frac{Q^2}{2M_{\mathrm{N}}^2}\frac{1}{x^2} = 2W_1 \tag{6.6.10b}$$

According to Eqs. (6.6.10a), (6.6.10b) and the definition of $x \equiv \frac{Q^2}{2M_N\nu}$, we can obtain further

$$F_1(x) \equiv 2M_\mathrm{N}W_1 = \sum_i f_i(x)Q_i^2 \tag{6.6.11}$$

$$F_2(x) \equiv \nu W_2 = x\sum_i f_i(x)Q_i^2 \tag{6.6.12}$$

Equations (6.6.11) and (6.6.12) show that it can indeed be derived from the parton model, the nucleon structure functions are only functions of the scaling variable x, rather than functions of two independent variables ν and Q^2. And in addition, the structure functions $F_1(x)$ and $F_2(x)$ satisfy

$$xF_1(x) = F_2(x) \tag{6.6.13}$$

Experiment shows that the data of the inelastic scattering of electrons to protons and neutrons conform to the relation of Eq. (6.6.13). So the experimental data indicate that the partons are particles with spin of 1/2.

6.7. The Unification of Parton and Quark Model

The experiment of electron–nucleon deep inelastic scattering shows that nucleon is composed of partons with spin of 1/2, while all hadrons including nucleon are consisted of quarks in quark model. Then what is the relation between these two models? Actually, the parton model and quark model can be consistent with each other. They can be unified into one model. Such unified model are called quark–parton model. In the quark model, nucleon is made up of three quarks: uud for proton, and udd for neutron, while the number of partons in nucleon is not restricted in the parton model. To unify these two models, we can think of the nucleon as something like this, that is, in addition to the quarks uud and/or udd in the nucleon that show its "property", it also includes a large number of fluctuating quark and antiquark pairs in nucleon: $\mathrm{u\bar{u}}$, $\mathrm{d\bar{d}}$, $\mathrm{s\bar{s}}$, $\mathrm{c\bar{c}}$, $\mathrm{b\bar{b}}$, etc. Among them, the quarks uud and udd that mark the nucleon's properties are called valence quarks, while the fluctuating quark–antiquark pairs are called sea quarks, because they are in the continuously fluctuating "sea". Sea quarks emerge due to quantum fluctuations. The existence of sea quarks can only be detected as high-energy electrons bombarding protons. It can be considered that the quark model is a low-energy image of the nucleon structure, while parton model is the high-energy image. QCD unifies these quantum images in such a way.

Define $f_{\mathrm{u}}(x)$, $f_{\mathrm{d}}(x)$, $f_{\mathrm{s}}(x)$, $f_{\mathrm{c}}(x)$ etc. as the probability density of the relevant quarks in the proton, while $f_{\bar{\mathrm{u}}}(x)$, $f_{\bar{\mathrm{d}}}(x)$, $f_{\bar{\mathrm{s}}}(x)$, $f_{\bar{\mathrm{c}}}(x)$ etc. are the probability density of antiquarks in the proton, where x is the fraction of the quarks's momentum relevant to the momentum of the proton in the infinity momentum frame. Note that these probability densities may not be normalized to 1. For example, the probability density for u quark in proton is normalized to 2, $\int_0^1 dxu(x) = 2$, because there are two valence quarks u in proton. Because the numbers of the quarks and antiquarks in the "sea" in nucleon are equal, the number of quarks minus the number of antiquarks should be equal to the number of valence quarks in the nucleon. Therefore, for proton there are the following results:

$$\int_0^1 [f_{\mathrm{u}}(x) - f_{\bar{\mathrm{u}}}(x)]\mathrm{d}x = 2 \tag{6.7.1}$$

$$\int_0^1 [f_{\mathrm{d}}(x) - f_{\bar{\mathrm{d}}}(x)]\mathrm{d}x = 1 \tag{6.7.2}$$

$$\int_0^1 [f_{\mathrm{s}}(x) - f_{\bar{\mathrm{s}}}(x)]\mathrm{d}x = 0 \tag{6.7.3}$$

$$\int_0^1 [f_{\mathrm{c}}(x) - f_{\bar{\mathrm{c}}}(x)]\mathrm{d}x = 0 \tag{6.7.4}$$

Because the charges of quarks are $Q_u = 2/3$, $Q_{\mathrm{d}} = -1/3$, $Q_{\mathrm{s}} = -1/3$, $Q_{\mathrm{c}} = 2/3$, we can obtain the structure function of proton according to Eq. (6.6.12)

$$\begin{aligned} F_2^{\mathrm{p}}(x) = x\bigg\{ & \left(\frac{2}{3}\right)^2 [f_{\mathrm{u}}(x) + f_{\bar{\mathrm{u}}}(x)] + \left(\frac{1}{3}\right)^2 [f_{\mathrm{d}}(x) + f_{\bar{\mathrm{d}}}(x)] \\ & + \left(\frac{1}{3}\right)^2 [f_{\mathrm{s}}(x) + f_{\bar{\mathrm{s}}}(x)] + \left(\frac{2}{3}\right)^2 [f_{\mathrm{c}}(x) + f_{\bar{\mathrm{c}}}(x)] + \cdots \bigg\} \end{aligned} \tag{6.7.5}$$

For neutrons, the distribution of u quarks is the same thing as d quark in proton, and the distribution of d quark is the same as that of u quark in proton. Therefore, making the substitution $f_{\mathrm{u}}(x) \leftrightarrow f_{\mathrm{d}}(x)$ and $f_{\bar{\mathrm{u}}}(x) \leftrightarrow f_{\bar{\mathrm{d}}}(x)$ in Eq. (6.7.5), one can get the structure function of neutron,

$$\begin{aligned} F_2^{\mathrm{n}}(x) = x\bigg\{ & \left(\frac{2}{3}\right)^2 [f_{\mathrm{d}}(x) + f_{\bar{\mathrm{d}}}(x)] + \left(\frac{1}{3}\right)^2 [f_{\mathrm{u}}(x) + f_{\bar{\mathrm{u}}}(x)] \\ & + \left(\frac{1}{3}\right)^2 [f_{\mathrm{s}}(x) + f_{\bar{\mathrm{s}}}(x)] + \left(\frac{2}{3}\right)^2 [f_{\mathrm{c}}(x) + f_{\bar{\mathrm{c}}}(x)] + \cdots \bigg\} \end{aligned} \tag{6.7.6}$$

where it has been assumed that the distributions of sea quarks are the same in both proton and neutron. The average momentum of some quarks in proton is

$$\begin{aligned} &P_{\mathrm{u}} = \int x\,[f_{\mathrm{u}}(x) + f_{\bar{\mathrm{u}}}(x)]\mathrm{d}x, \quad P_{\mathrm{d}} = \int x\,[f_{\mathrm{d}}(x) + f_{\bar{\mathrm{d}}}(x)]\mathrm{d}x \\ &P_{\mathrm{s}} = \int x\,[f_{\mathrm{s}}(x) + f_{\bar{\mathrm{s}}}(x)]\mathrm{d}x, P_{\mathrm{c}} = \int x\,[f_{\mathrm{c}}(x) + f_{\bar{\mathrm{c}}}(x)]\mathrm{d}x \\ &\qquad\vdots \end{aligned} \tag{6.7.7}$$

Using Eqs. (6.7.7), (6.7.5) and (6.7.6), we can obtain the following relations:

$$\int F_2^p(x)\mathrm{d}x = \frac{4}{9}P_{\mathrm{u}} + \frac{1}{9}P_{\mathrm{d}} + \frac{1}{9}P_{\mathrm{s}} + \frac{4}{9}P_{\mathrm{c}} + \cdots \tag{6.7.8}$$

$$\int F_2^n(x)\mathrm{d}x = \frac{4}{9}P_{\mathrm{d}} + \frac{1}{9}P_{\mathrm{u}} + \frac{1}{9}P_{\mathrm{s}} + \frac{4}{9}P_{\mathrm{c}} + \cdots \tag{6.7.9}$$

It has been experimentally determined by high-energy neutrino (antineutrino) scattering experiments [30]

$$\frac{\int [f_{\bar{\mathrm{u}}}(x) + f_{\bar{\mathrm{d}}}(x) + f_{\bar{\mathrm{s}}}(x) + f_{\bar{\mathrm{c}}}(x)]\,\mathrm{d}x}{\int [f_{\mathrm{u}}(x) + f_{\bar{\mathrm{u}}}(x) + \cdots + f_{\mathrm{c}}(x) + f_{\bar{\mathrm{c}}}(x) + \cdots]\,\mathrm{d}x} \approx 0.15 \tag{6.7.10}$$

This result indicates that the contribution of sea quarks is not completely negligible. Another experimental result is

$$\frac{\int f_{\bar{\mathrm{s}}}(x)\mathrm{d}x}{\int [f_{\mathrm{u}}(x) + f_{\bar{\mathrm{u}}}(x) + \cdots + f_{\mathrm{c}}(x) + f_{\bar{\mathrm{c}}}(x) + \cdots]\,\mathrm{d}x} \approx 0.025 \tag{6.7.11}$$

The result shows that the probability of strange quark emerging in nucleons is at the order of a few percent. Since the masses of quarks $\mathrm{c}, \bar{\mathrm{c}}$ and $\mathrm{b}, \bar{\mathrm{b}}$ are heavier, it is reasonable to assume that these quarks emerge in nucleon with even lower probability. Neglect the contribution of s quark, c quark and heavier quarks in Eqs. (6.7.8) and (6.7.9), then according to experimental result

$$\int F_2^p(x)\mathrm{d}x \approx 0.18 \quad \text{and} \quad \int F_2^n(x)\mathrm{d}x \approx 0.12 \tag{6.7.12}$$

one can obtain

$$P_{\mathrm{u}} \approx 0.36, \quad P_{\mathrm{d}} \approx 0.18 \tag{6.7.13}$$

This result suggests that the momenta carried by quark and antiquarks in nucleon are only about half of the momentum of the nucleon, so there should be other constituent particles in the nucleons. According to QCD (quantum chromodynamics), the quarks in the nucleons are bound together by color field. The quantum of color field is gluon, which is vector with spin of 1, and not involved in electromagnetic interaction. Therefore, in deep inelastic electron–nucleon scattering experiment, it is not possible to detect the existence of gluon by electromagnetic interaction. So one interpretation to the experimental result is that the other half of the momentum of the nucleon is carried by gluons.

References

1. Tomonaga S. *Prog. Theor. Phys.*, 1946, 1: 27.
2. Tomonaga S, Oppenheimer J R. *Phys. Rev.*, 1948, 74: 224.
3. Feynman R. *Phys. Rev.*, 1948, 74: 1430.
4. Feynman R. *Phys. Rev.*, 1949, 76: 749.
5. Feynman R. *Phys. Rev.*, 1949, 76: 769.
6. Schwinger J S. *Phys. Rev.*, 1948, 73: 416.
7. Schwinger J S. *Phys. Rev.*, 1948, 74: 1439.
8. Schwinger J S. *Phys. Rev.*, 1948, 75: 651.
9. Schwinger J S. *Phys. Rev.*, 1949, 76: 790.
10. Lamb W E, Retherford R C. *Phys. Rev.*, 1947, 72: 241.
11. Foley H M, Kusch P. *Phys. Rev.*, 1948, 73: 412.
12. Coffin T, Garwin R L, Penman S, *et al. Phys. Rev.*, 1958, 109: 973.
13. Bethe H A. *Phys. Rev.*, 1947, 72: 339.
14. Kroll N M, Lamb W E. *Phys. Rev.*, 1949, 75: 388.
15. French J B, Weisskopf V F. *Phys. Rev.*, 1949, 75: 1240.
16. Barber W C, Gittelman B, O'Neill G K, *et al. Phys. Rev.*, 1971, D3: 2796.
17. Adeva B, *et al. Phys. Rev. Lett.*, 1982, 48: 721.
18. Adeva B, *et al.* (MARK-J Collab.) *Progr. Phys.*, 1984, 4: 301.
19. ALEPH and DELPHI and L3 and OPAL and LEP Electroweak Group and SLD Heavy Flavor Group Collab. e-print hep-ex/0212036.
20. Chambers E E, Hofstadter R. *Phys. Rev.*, 1956, 103: 1454.
21. Bloom E, *et al. Report to Fourteenth International Conference on High Energy Physics* Vienna, 1968 (unpublished).
22. Bloom E D, *et al. Phys. Rev. Lett.*, 1969, 23: 930.
23. Breidenbach M, *et al. Phys. Rev. Lett.*, 1969, 23: 935.
24. Bjorken J D. *Phys. Rev.*, 1969, 179: 1547.
25. Feynman R P. *Phys. Rev. Lett.*, 1969, 23: 1415.
26. Rosenbluth M N. *Phys. Rev.*, 1950, 79: 615.
27. Ernst F J, Sachs R G, Wali K C. *Phys. Rev.*, 1960, 119: 1105.
28. Sachs R G. *Phys. Rev.*, 1961, 126: 2256.
29. Hughes E B, Griffy T A, Yearian M R, *et al. Phys. Rev.*, 1965, 139: B458.
30. de Groot J G H, *et al. Z. Phys.*, 1979, C1: 143.

This page intentionally left blank

Chapter 7

Weak Interactions

7.1. Looking Back to History

The existence of the weak interactions was firstly discovered in β decay of nuclei in late 19th century. Later, in order to explain the continuity of the β spectrum Pauli proposed the neutrino assumption in his letter to L. Meitner in 1930, but the formal paper was published in 1933 [1].

In early 20th century, it was found through investigation to large amount of nuclear β decays that β decays basically have two kinds: one is Fermi type [2], the nuclear spins for the initial and final states of the decay are all zero, and the angular momenta carried away by the final state neutrino and electron are also zero, that is so-called $0 \to 0$ transition, e.g.

$$\mathrm{O}^{14} \to \mathrm{N}^{14*} + \mathrm{e}^{+} + \nu_{\mathrm{e}} \tag{7.1.1}$$

Another is Gamov–Teller type [3], the nuclear spin of the initial state is zero, that of the final state is 1, the total momentum carried away by neutrino and electron is 1, for instance

$$\mathrm{He}^{6} \to \mathrm{Li}^{6} + \mathrm{e}^{-} + \bar{\nu}_{\mathrm{e}} \tag{7.1.2}$$

Fermi assumed that the basic process for nuclear β decay is

$$\mathrm{n} \to \mathrm{p} + \mathrm{e}^{-} + \bar{\nu}_{\mathrm{e}} \tag{7.1.3}$$

while for β^{+} decay is

$$\mathrm{p} \to \mathrm{n} + \mathrm{e}^{+} + \nu_{\mathrm{e}} \tag{7.1.4}$$

Of course, because $m_{\mathrm{n}} > m_{\mathrm{p}}$, the process in Eq. (7.1.4) can happen only for the proton bounded in nuclei. Fermi proposed [2] further a theory of four fermion current–current interactions for describing β decay. Its effective Hamiltonian is

$$H = \frac{G}{\sqrt{2}} \left(\bar{\psi}_{\mathrm{p}} \gamma_\mu \psi_{\mathrm{n}}\right) \left(\bar{\psi}_{\mathrm{e}} \gamma^\mu \psi_\nu\right) \tag{7.1.5}$$

where, $J_\mu = \bar{\psi}_{\mathrm{p}} \gamma_\mu \psi_{\mathrm{n}}$ is hadronic current, $j_\mu = \bar{\psi}_{\mathrm{e}} \gamma_\mu \psi_\nu$ is leptonic current.

Equation (7.1.5) can explain Fermi-type β decay, but not Gamov–Teller one. To explain later needs to introduce axial current

$$J_\mu = \bar{\psi}_{\mathrm{p}} \gamma_\mu \gamma_5 \psi_{\mathrm{n}}$$

In 1956 T.D. Lee and C.N. Yang firstly proposed that parity may be not conserved in weak interaction processes [4]. This assumption is verified by experiments in 1957 [5]. This pushed Feynman and Gell-Mann to propose in 1958 the universal V–A theory for weak interactions [6]. Its essential idea is: for five different fermionic bi–linear couplings $\bar{\psi}_A \Gamma_i \psi_B$ (Γ_i denote the five different combinations of γ matrices S, P, T, V, A) only the V–A coupling is effective. The proposal of V–A theory put the calculations of weak interactions on a solid theory foundation.

In order to use one coupling constant G_{F} and V–A theory to explain all weak interaction processes including $|\Delta S| = 0$ and $|\Delta S| = 1$ processes, Cabibbo [7] introduced mixing of hadronic currents in 1963

$$J_\mu = J_\mu^{\Delta S=0} \cos\theta_{\mathrm{C}} + J_\mu^{\Delta S=1} \sin\theta_{\mathrm{C}} \tag{7.1.6}$$

θ_{C} is called Cabibbo angle. At quark level, Eq. (7.1.6) can be written as

$$J_\mu = \bar{u}\gamma_\mu \left(1-\gamma_5\right) \left(\mathrm{d}\cos\theta_{\mathrm{C}} + \mathrm{s}\sin\theta_{\mathrm{C}}\right) \tag{7.1.7}$$

From Eq. (7.1.7) we see that d and s quarks participate the weak interactions only through the mixing state $\mathrm{d}\cos\theta_{\mathrm{C}} + \mathrm{s}\sin\theta_{\mathrm{C}}$. The mixing matrix is 2×2 matrix. This kind of mixing causes flavor-changing neutral current, for example, the terms like:

$$\bar{\mathrm{d}}\gamma_\mu \left(1-\gamma_5\right) \mathrm{s}, \quad \bar{\mathrm{s}}\gamma_\mu \left(1-\gamma_5\right) \mathrm{d} \tag{7.1.8}$$

The existence of these terms contradicts with experiments.

In order to solve this problem, in 1970, S. Glashow, J. Iliopoulos and L. Maiani proposed GIM mechanism [8]. The main idea is to introduce the fourth flavored quark c (charm) to join weak interactions. This mechanism can eliminate flavor-changing neutral current like the term in Eq. (7.1.8). In 1973, Kobayashi and Maskawa [9] extended flavor mixing to the case of three generations of quarks (u, d; c, s; t, b), thus introducing 3×3 quark mixing matrix which is called CKM (Cabibbo–Kobayashi–Maskawa) matrix. The extension from $2 \times 2 \to 3 \times 3$ mixing matrix induced an essential change, that is the appearance of a complex phase in 3×3 mixing matrix which can explain CP violation.

In 1964 CP violation was observed experimentally in the decay of $K_{\mathrm{L}} \to 2\pi$ [10]. This phenomenon can be explained by the complex phase in the CKM matrix.

Although the effective (V–A) theory of four fermion point-like interactions achieved a greet success, but when the center-of-mass energy exceeds 300 GeV, the cross section for $\nu_\mu + \mathrm{e}^- \to \mu^- + \nu_\mathrm{e}$ scattering process will violate the conservation of probability. In order to overcome this difficulty, Lee, Rosenbluth and Yang [11] assumed that the current–current weak interactions are induced by exchanging the massive intermediate vector boson W. This idea comes from the inspiration of the electromagnetic current–current interactions which is induced by exchanging photons. Because the weak force is short ranged, so the W boson must be very heavy. The intermediate vector boson assumption solved the difficulty of four fermion point-like interactions at high-energy limit, but this theory is not renormalizable. This non-renormalizability problem is solved until the emerge of the standard model of the unified electroweak interaction [12]. This standard model will be discussed in detail in Chapter 8.

7.2. Classification of Weak Decays

The strength of weak interactions is much weaker than that of electromagnetic and strong interactions. Why can we still observe the effects of weak interaction processes? The reason is that for these processes the strong and electromagnetic interactions are forbidden by conservation laws including electric charge and lepton number conservation, etc.

We find from experiments that $(\nu_\mathrm{e}, \mathrm{e}^-)$ carry lepton number $L_\mathrm{e} = 1$, (ν_μ, μ^-) carry lepton number $L_\mu = 1$, and (ν_τ, τ^-) carry lepton number $L_\tau = 1$. All their anti-particles carry lepton numbers L_e, L_μ, L_τ as -1, respectively. $L_{\mathrm{e},\, \mu,\, \tau}$ conserve in all weak, electromagnetic, and strong interaction processes, and further more L_e, L_μ, L_τ conserve separately. These separate conservations hold only when the masses of ν_e, ν_μ, ν_τ are accurately zero. Till now, the masses of neutrinos measured in experiments are all very small [13]:

$$\begin{aligned} &m_{\bar{\nu}_{\mathrm{e}}} < 2\,\mathrm{eV} \\ &m_{\nu_\mu} < 0.19\,\mathrm{MeV} \\ &m_{\nu_\tau} < 18.2\,\mathrm{MeV} \end{aligned} \tag{7.2.1}$$

while the astrophysical observations gives a constrain for the sum of masses of all kinds of neutrinos, i.e. [14]

$$\sum_i m_{\nu_i} < 0.28\,\mathrm{eV} \tag{7.2.2}$$

So when discussing most of the processes of weak interactions (except for the neutrino oscillations) the assumption for $m_{\nu_{\mathrm{e}}} = m_{\nu_\mu} = m_{\nu_\tau} = 0$ is very good approximation.

Neutrino oscillation has been already verified in experiments, for example, we have already observed the atmospheric neutrino transformation of $\nu_\mu \to \nu_\tau$, the solar neutrino transformation of $\nu_{\mathrm{e}} \to \nu_\mu, \nu_\tau$ etc., so neutrino masses cannot be zero definitely. Thus the separate conservation of L_{e}, L_μ, L_τ must be broken in neutrino oscillation experiments, but this conservation law still holds in other experiments. We shall come back to these problems in Chapter 10 later.

Weak decay processes can be classified into three classes according to the particle type of participation

(1) **Pure leptonic decays:** In these weak decay processes, there are only leptons in the initial and final states, e.g.

$$\begin{aligned} &\tau^- \to \mu^- + \bar{\nu}_\mu + \nu_\tau \\ &\mu^- \to \mathrm{e}^- + \bar{\nu}_{\mathrm{e}} + \nu_\mu \\ &\nu_{\mathrm{e}} + \mathrm{e}^- \to \nu_{\mathrm{e}} + \mathrm{e}^- \end{aligned}$$

Obviously, L_{e}, L_μ, L_τ conserve separately in the above processes.

(2) **Semi-leptonic decays:** In these decay processes, there are both leptons and hadrons in the initial and final states, e.g.

$$\begin{aligned} &\mathrm{n} \to \mathrm{p} + \mathrm{e}^- + \bar{\nu}_{\mathrm{e}}, \quad \mathrm{p} + \bar{\nu}_{\mathrm{e}} \to \mathrm{n} + \mathrm{e}^+ \\ &\pi^\pm \to \mu^\pm + \nu_\mu(\bar{\nu}_\mu), \quad \mathrm{K}^\pm \to \mu^\pm + \nu_\mu(\bar{\nu}_\mu) \\ &\pi^\pm \to \mathrm{e}^\pm + \nu_{\mathrm{e}}(\bar{\nu}_{\mathrm{e}}), \quad \mathrm{K}^+ \to \pi^0 + \mathrm{e}^+ + \nu_{\mathrm{e}} \end{aligned}$$

etc. Obviously, in the above processes, the lepton numbers L_{e}, L_μ conserve separately.

(3) **Non-leptonic decays:** In these processes, only hadrons participate, for instance,

$$\begin{aligned} \Lambda &\to \mathrm{p} + \pi^-, \\ &\to \mathrm{n} + \pi^0, \\ \mathrm{K}^+ &\to \pi^+ + \pi^0, \\ &\to \pi^+ + \pi^- + \pi^+ \end{aligned}$$

The reason for putting $\Lambda\mathrm{K}^+$ decays in the above processes into the category of weak interaction processes is that the decay times of these processes are very long, i.e. longer than that of strong interaction processes for several orders of magnitude.

7.3. Nuclear β Decay

We have pointed out in Sec. 7.1 that nuclear β decays are divided to Fermi type and Gamov–Teller type. Take the nucleus, the electron, and the neutrino as a whole system, then the Hamiltonian of the system is

$$H = H_{\mathrm{N}} + H_{\mathrm{e}} + H_\nu + H_{\mathrm{W}} \tag{7.3.1}$$

where H_{N} is the nuclear Hamiltonian, H_{e} is the electron Hamiltonian, and H_ν is the neutrino Hamiltonian, and H_{W} is the weak interaction Hamiltonian which in general can be written as

$$H_{\mathrm{W}} = \frac{G}{\sqrt{2}} \bar{\psi}_{\mathrm{p}} \gamma^\mu (a - b\gamma_5) \psi_{\mathrm{n}} \bar{\psi}_{\mathrm{e}} \gamma_\mu (1 - \gamma_5) \psi_{\nu_{\mathrm{e}}} \tag{7.3.2}$$

where term a only contributes to the Fermi type of β decays, and term b only contributes to Gamov–Teller type of β decays. The magnitude of a, b and their deviation from 1 indicate the amount of the strong interaction corrections. From experiments, we have $a \approx 1$, $b \approx 1.25$, the coupling constant G is actually

$$G = G_{\mathrm{F}} \cos\theta_{\mathrm{C}} \tag{7.3.3}$$

where θ_{C} is the Cabibbo angle and G_{F} is the universal weak interaction constant

$$G_{\mathrm{F}} = 1.166 \times 10^{-5}\,\mathrm{GeV}^{-2}, \quad \cos\theta_{\mathrm{C}} \approx 0.974 \tag{7.3.4}$$

The transition probability of β decay is

$$W = \frac{2\pi}{\hbar}|\langle f|H_{\mathrm{W}}|\mathrm{i}\rangle|^2\rho \tag{7.3.5}$$

where $\rho = \mathrm{d}N/\mathrm{d}E$ is the final state density in unit electron energy spacing. The calculation for the probability in Eq. (7.3.5) involves the integration of the nuclear wave functions for the initial and final states, so is very complicated. We shall not give it here. For readers who are interested in it can refer any text books of nuclear theory [15].

The calculation gives that the final state density N for β decay satisfies

$$N(p)\mathrm{d}p \propto p^2(E_0 - E)^2\mathrm{d}p \tag{7.3.6}$$

where p is the momentum of the final state electron, here we assume $m_\nu = 0$, and neglect the correction of the nuclear Coulomb field when the electron flying out through the nucleus, E_0 is the maximum energy of the final state electron. If we consider $m_\nu \neq 0$ and the nuclear Coulomb field corrections, then Eq. (7.3.6)becomes

$$N(p)F(Z,p)\mathrm{d}p \propto p^2(E_0 - E)^2\sqrt{1 - \left(\frac{m_\nu}{E_0 - E}\right)^2}\,\mathrm{d}p \tag{7.3.7}$$

where $F(Z,p)$ is the Coulomb correction function which can be calculated theoretically. The larger the nuclear charge Z is, the bigger the Coulomb correction function $F(Z,p)$ becomes. Figure 7.1 shows the relation for $\sqrt{N(p)F(Z,p)/p^2}$ and the electron energy E. If $m_\nu = 0$, then the relation becomes a straight line. Its intersection point with the horizontal axes is the maximum electron energy E_0 (as shown by the dashed line). If $m_\nu \neq 0$, the end part of the straight line will be curved. The maximum

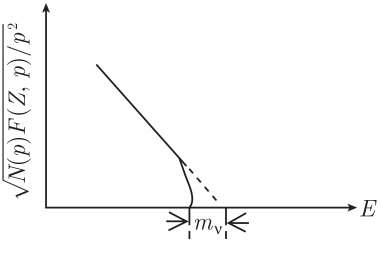


Fig. 7.1. The relation of $\sqrt{NF/p^2}$ and E (Kurie plot).

energy of the electron will be reduced by amount of m_ν. Figure 7.1 is called Kurie plot.

In experiments, some people use the Kurie plot for Tritium β decay to measure the mass of the electron neutrino. But in practice the data analysis is more complicated than using Kurie plot simply. In experiments, people use Tritium molecule T_2 instead of Tritium atom, and they need to measure the differential decay rate $\mathrm{d}R(E)/\mathrm{d}E$ and the relation of the integrated decay rate $R(E)$ with the neutrino mass squared (m_ν^2) etc. At present, the upper bound of electron neutrino mass for this kind of experiments is 2 eV, i.e.

$$m_{\nu_\mathrm{e}} < 2\,\mathrm{eV} \tag{7.3.8}$$

For detail, please refer [16] and references therein.

7.4. The Discovery of Parity Violation

7.4.1. *τ–θ Puzzle*

In 1947 many new strange particles were observed, among them, two particles τ^+ and θ^+ showed very strange property: their masses are almost the same, but their decay final states have different parity:

$$\theta^+ \to \pi^+\pi^0 \quad \text{(the parity of the final state is positive)}$$
$$\tau^+ \to \pi^+\pi^-\pi^+ \quad \text{(the parity of the final state is negative)}$$

At that time scientists thought in general parity was conserved in weak interaction processes, so τ^+ and θ^+ seemed not to be the same particle. But as they have the same mass, τ^+ and θ^+ seemed to be the same particle. That is, the so-called τ–θ puzzle.

In 1956, T.D. Lee and C.N. Yang analyzed [4] all kind of experiments at that time and found no evidence which indicated parity conservation in weak interaction processes. They gave a bold assumption that parity does not conserve in weak interaction processes, and suggested a series of experiments to test parity non-conservation. If parity is not conserved in weak interaction processes, then taking τ^+, θ^+ as one particle will be no problem. The τ–θ puzzle can be easily solved. Now we know that τ^+ and θ^+ are just the same particle — K^+ meson.

Soon after, parity violation was verified in the cobalt $60(\mathrm{Co}^{60})\beta$ decay and the weak decays of π mesons and μ [5].

7.4.2. *Parity violation in β decays of cobalt 60 nuclei*

The process for the β decay of cobalt 60 is

$$\mathrm{Co}^{60} \to \mathrm{Ni}^{60} + \mathrm{e}^- + \bar{\nu}_\mathrm{e} \tag{7.4.1}$$

Its basic process is $\mathrm{n} \to \mathrm{p} + \mathrm{e}^- + \bar{\nu}_\mathrm{e}$.

In this experiment, the sample of Co^{60} is put in extremely low temperature (0.01 K), switching on the magnetic field to make the Co^{60} nuclear polarizations along the magnetic field direction. Then measure the counting of the outgoing electrons at the angles θ and $180° - \theta$ as shown in Fig. 7.2.

In experiments, the measurements sketched in Fig. 7.2 are carried out through changing the direction of the magnetic field ($\vec{B} \to -\vec{B}$) so changing the nucleus Co^{60} polarization to count for $180° - \theta$. The measurements show that the count number of electrons parallel to the nucleus spin $\vec{\sigma}$ (i.e. $\vec{B}$) is much less than that of antiparallel to $\vec{\sigma}$, the difference is about 40%. Put the momentum of the electron as $\vec{p}$, take $\langle \vec{\sigma} \cdot \vec{p} \rangle$ to denote the average value of the measured $\vec{\sigma} \cdot \vec{p}$. Under space reflection, $\vec{p} \to -\vec{p}$, $\vec{\sigma} \to \vec{\sigma}$, so

$$\langle \vec{\sigma} \cdot \vec{p} \rangle \to -\langle \vec{\sigma} \cdot \vec{p} \rangle \tag{7.4.2}$$

If the parity is conserved, then the measured results should be

$$\langle \vec{\sigma} \cdot \vec{p} \rangle = -\langle \vec{\sigma} \cdot \vec{p} \rangle = 0 \tag{7.4.3}$$

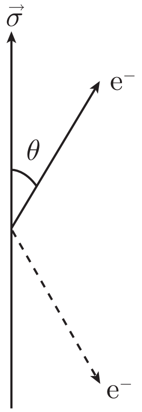


Fig. 7.2. Sketch of the Co^{60} experiment.

But in the above experiment of Co^{60} the outgoing electrons prefer to choose the direction antiparallel to the nucleus spin, so we have

$$\langle\vec{\sigma}\cdot\vec{p}\rangle_{\mathrm{Co}^{60}} < 0 \tag{7.4.4}$$

Comparing with Eq. (7.4.3), the parity conservation in the $\mathrm{Co}^{60}\beta$ decay is broken.

For detail, please refer the paper by C.S.Wu's group [5].

7.5. The V–A Theory of Weak Interactions

The four Fermion current–current interaction theory proposed by Fermi is a vector current coupling theory, it can only explain the Fermi type β decay. In order to explain the Gamov–Teller type β decay we need to introduce the pseudovector current coupling. In general current–current couplings have five different forms, i.e.

$$H = \sum_i \frac{G_i}{\sqrt{2}}\bar{\psi}_A\Gamma_i\bar{\psi}_B\bar{\psi}_C\Gamma_i\psi_D \tag{7.5.1}$$

where Γ_i take five different γ matrix combinations:

$$\begin{aligned}
&(1)\quad \text{scalar type S}, \Gamma_i = 1;\\
&(2)\quad \text{peudoscalar type P}, \Gamma_i = \gamma_5;\\
&(3)\quad \text{tensor type T}, \Gamma_i = \sigma_{\mu\nu} = \tfrac{\mathrm{i}}{2}\left(\gamma_\mu\gamma_\nu - \gamma_\nu\gamma_\nu\right);\\
&(4)\quad \text{vector type V}, \Gamma_i = \gamma_\mu;\\
&(5)\quad \text{pseudovector type A}, \Gamma_i = \gamma_\mu\gamma_5.
\end{aligned} \tag{7.5.2}$$

In 1958, after analyzed large amount of weak interaction experiments, Feynman and Gell-Mann proposed a universal V–A weak interaction theory [6], i.e. in the S, P, T, V, A five currents, only V and A currents are relative, and appear with V–A combination. We now explain the application of the V–A theory in different kinds of weak decays below.

7.5.1. *Pure leptonic decay*

We first look at the decay of $\mu^- \to \mathrm{e}^- + \bar{\nu}_\mathrm{e} + \nu_\mu$. According to the V–A theory, the effective Hamiltonian of this process can be written as

$$H_\mathrm{W} = \frac{G_\mu}{\sqrt{2}}\bar{\mathrm{e}}\gamma_\mu\left(1-\gamma_5\right)\nu_\mathrm{e}\bar{\nu}_\mu\gamma^\mu\left(1-\gamma_5\right)\mu + \mathrm{h.c.} \tag{7.5.3}$$

In Eq. (7.5.3) the symbols of particles represent the corresponding fields, i.e. $\bar{\mathrm{e}} = \bar{\psi}_{\mathrm{e}}$ etc. G_μ is the coupling constant. It is easy to calculate the μ decay width Γ_μ as

$$\Gamma_\mu = \frac{1}{\tau_\mu} \approx \frac{G_\mu^2 m_\mu^5}{192\pi^3} \tag{7.5.4}$$

where the terms of the order of m_{e}^2/m_μ^2 are neglected.

Substituting the measured μ lifetime τ_μ into Eq. (7.5.4) we obtain the value of G_μ which is just equal the universal Fermi weak coupling constant G_{F} (see Eq. (7.3.4)). For $\tau^- \to \mu^- + \bar{\nu}_\mu + \nu_\tau$, the effective Hamiltonian can be written as

$$H_{\mathrm{W}} = \frac{G_{\tau_\mu}}{\sqrt{2}} \bar{\mu}\gamma_\mu(1-\gamma_5)\nu_\mu \bar{\nu}_\tau \gamma^\mu (1-\gamma_5)\,\tau \tag{7.5.5}$$

For $\tau^- \to \mathrm{e}^- + \bar{\nu}_{\mathrm{e}} + \nu_\tau$, the effective Hamiltonian is

$$H_{\mathrm{W}} = \frac{G_{\tau_{\mathrm{e}}}}{\sqrt{2}} \bar{\mathrm{e}}\gamma_\mu(1-\gamma_5)\nu_{\mathrm{e}} \bar{\nu}_\tau \gamma^\mu (1-\gamma_5)\,\tau \tag{7.5.6}$$

From Eqs. (7.5.5) and Eq. (7.5.6), we can compute the corresponding branching ratios

$$\mathrm{Br}\left(\tau^- \to \mu^- + \bar{\nu}_\mu + \nu_\tau\right) = \frac{1}{\Gamma_\tau} \frac{G_{\tau_\mu}^2 m_\tau^5}{192\pi^3} \tag{7.5.7}$$

$$\mathrm{Br}\left(\tau^- \to \mathrm{e}^- + \bar{\nu}_{\mathrm{e}} + \nu_\tau\right) = \frac{1}{\Gamma_\tau} \frac{G_{\tau_{\mathrm{e}}}^2 m_\tau^5}{192\pi^3} \tag{7.5.8}$$

where Γ_τ is the total decay width of τ^- and $\frac{1}{\Gamma_\tau} = \tau_\tau$ is the τ^- lifetime. Using the measured value

$$\tau_\tau = (290.6 \pm 1.0) \times 10^{-15}\mathrm{s}$$

from Eqs. (7.5.7) and Eq. (7.5.8) we can calculate the coupling constants

$$G_{\tau_\mu} \approx G_{\tau_{\mathrm{e}}} \approx G_{\mathrm{F}} \tag{7.5.9}$$

This shows that pure leptonic decays have a universal weak interaction coupling constant G_{F}.

In 1992, the Particle Data Group (PDG) published the value of the τ mass [17]

$$m_\tau = 1784.1\,\begin{matrix}+2.7\\-3.6\end{matrix}\,\text{MeV} \tag{7.5.10}$$

So we can deduce the ratio of the coupling constants for the τ decays of $\mathrm{e}\bar{\nu}_\mathrm{e}\nu_\tau$ and $\mu\bar{\nu}_\mu\nu_\tau$

$$(G_{\tau_\mathrm{e}}/G_{\tau_\mu})^2 = 0.941 \pm 0.025 \tag{7.5.11}$$

This value deviates 2.4σ from that of the lepton universality of Eq. (7.5.9).

In 1992, BES Collaboration remeasured the τ mass at the threshold of the process $\mathrm{e}^+\mathrm{e}^- \to \tau^+\tau^-$ [18]. The measured value is

$$m_\tau = 1776.9^{+0.4}_{-0.5} \pm 0.2\,\text{MeV} \tag{7.5.12}$$

This makes the establishment of the lepton universality of Eq. (7.5.9). This is one of the important contributions of BES Collaboration.

7.5.2. *Semi-leptonic decay*

In Sec. 7.3, we have discussed the nucleus β decay. Now we discuss π, K meson decays, and introduce the concept of helicity suppression.

We first discuss π meson two-body leptonic decays, i.e. so-called π_{l2}:

$$\pi^\pm \to l^\pm + \nu_l(\bar{\nu}_l) \tag{7.5.13}$$

where $l^\pm = \mu^\pm$ or $\mathrm{e}^\pm$.

In the above π_{l2} decays $\Delta S = 0$. According to Cabibbo theory (see Sec. 7.6) the effective Hamiltonian is

$$H_\mathrm{W} = \frac{G_\mathrm{F}}{\sqrt{2}}\cos\theta_\mathrm{C}(J^+_\mu j^{-\mu} + \text{h.c.}) \tag{7.5.14}$$

where the leptonic current

$$j^-_\mu = \bar{l}\gamma^\mu(1-\gamma_5)\nu_l \tag{7.5.15}$$

while the hadronic current

$$J^+_\mu = V^+_\mu - A^+_\mu \tag{7.5.16}$$

i.e. vector current V_μ minus axial vector current A_μ (V–A theory). We will give their concrete form at the quark level in Sec. 7.6.

The matrix element of Eq. (.5.13) is

$$\left\langle l^{\mp}\nu\right| H_{\mathrm{W}} \left|\pi^{\mp}\right\rangle = \frac{G_{\mathrm{F}}\cos\theta_{\mathrm{C}}}{\sqrt{2}} \left\langle l^{\mp}\nu\right| j^{\mp\mu} \left|0\right\rangle \left\langle 0\right| J_{\mu}^{\pm} \left|\pi^{\mp}\right\rangle \tag{7.5.17}$$

Whenever we know $\langle 0| J_\mu |\pi\rangle$ we can easily calculate the decay rate (decay width) from Eq. (7.5.17). According to Eq. (7.5.16),

$$\langle 0| J_\mu |\pi\rangle = \langle 0| V_\mu |\pi\rangle - \langle 0| A_\mu |\pi\rangle \tag{7.5.18}$$

Because $|\pi\rangle$ is a pseudoscalar, V_μ is a vector, so at space inversion $\langle 0| V_\mu(0) |\pi\rangle$ should be axial vector. But π meson has no spin, $\langle 0| V_\mu(0) |\pi\rangle$ can construct an axial vector only by π meson momentum k_μ, but this is not possible, because k_μ is a vector. Thus we only can have

$$\langle 0| V_\mu(0) |\pi\rangle = 0 \tag{7.5.19}$$

But $\langle 0| A_\mu(0) |\pi\rangle$ is a vector. It can be constructed only by k_μ. Using translation invariance of operator, we obtain

$$\langle 0| A_\mu^{\pm}(x) \left|\pi^{\mp}\right\rangle = \langle 0| \mathrm{e}^{\mathrm{i}\hat{P}\cdot x} A_\mu^{\pm}(0) \mathrm{e}^{-\mathrm{i}\hat{P}\cdot x} \left|\pi^{\mp}\right\rangle = \mathrm{i}F_\pi k_\mu \mathrm{e}^{-\mathrm{i}k\cdot x} \tag{7.5.20}$$

where F_π is a constant which is called π meson decay constant.

It is easy to calculate out the matrix element $\langle \mathrm{e}\nu| j_\mu |0\rangle$, so we can compute the decay width

$$\Gamma(\pi_{l2}) = \frac{G_{\mathrm{F}}^2 \cos^2\theta_{\mathrm{C}} F_\pi^2 m_l^2}{8\pi m_\pi^3} (m_\pi^2 - m_l^2)^2 \tag{7.5.21}$$

Because the branching ratios for other decay channels $\leq 10^{-4}$, $\pi^{\pm}$ lifetime is determined mainly by $\pi \to \mu\nu$ decay. Substituting μ mass into Eq. (7.5.21) we have

$$\frac{1}{\tau_\pi} = \Gamma(\pi_{l2}) = \frac{G_{\mathrm{F}}^2 \cos^2\theta_{\mathrm{C}} F_\pi^2 m_\mu^2}{8\pi m_\pi^3} (m_\pi^2 - m_\mu^2)^2 \tag{7.5.22}$$

Substituting the measured tau lifetime $\tau_\pi = (2.6033 \pm 0.0005) \times 10^{-8}$s into the above equation (here we use the natural unit system $\hbar = c = 1$) we get

$$F_\pi \approx 132\,\mathrm{MeV} \tag{7.5.23}$$

Note that in Eq. (7.5.21)$\Gamma(\pi_{l2}) \propto m_l^2$. Because $m_{\mathrm{e}}^2 \ll m_\mu^2$, so the decay rate of $\pi \to \mathrm{e}\nu$ is very small. If $m_l = 0$, then the decay $\pi \to l\nu$ is forbidden,

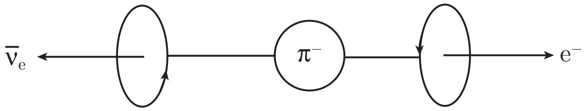


Fig. 7.3. Sketch of the decay $\pi^- \to e + \bar{\nu}_e$.

that is so-called helicity suppression. For example, consider the decay

$$\pi^- \to e + \bar{\nu}_e \tag{7.5.24}$$

In Fig. 7.3 we sketch the decay picture in π^- rest system. Because π^- is a pseudo-scalar particle with no spin, while $\bar{\nu}_e$ has right-handed helicity, in order to keep momentum conservation, the electron must has right-handed helicity. If $m_e = 0$, then the electron has only left-handed helicity.

When $m_e \neq 0$, at π meson rest frame, the probability for the electron to have right-handed helicity is

$$1 - v_e = 1 - \frac{p_e}{E_e} = \frac{2m_e^2}{m_\pi^2 + m_e^2} \tag{7.5.25}$$

This is just the reason for Eq. (7.5.21) to have the factor of m_l^2, and this is also the meaning of the helicity suppression.

From Eq. (7.5.21)

$$\frac{\Gamma(\pi_{e2})}{\Gamma(\pi_{\mu 2})} = \frac{m_e^2(m_\pi^2 - m_e^2)^2}{m_\mu^2(m_\pi^2 - m_\mu^2)^2} \sim 10^{-4} \tag{7.5.26}$$

so we can see that the branching ratio for the decay $\pi \to e\nu$ is indeed very small.

Similarly, we can compute the decay widths of $K \to l\nu$ decays. Because for these decays $\Delta S = 1$, so the factor of $\sin\theta_C$ will appear (see Sec. 7.6)

$$\Gamma(K_{l2}) = \frac{G_F^2 f_K^2 m_l^2}{8\pi m_K^3}(m_K^2 - m_l^2)^2 \sin^2\theta_C \tag{7.5.27}$$

For non-leptonic decay, we shall discuss it later in Sec. 7.10.

7.6. Cabibbo Theory and GIM Mechanism

In 1963, Cabibbo [7] firstly introduced the hadronic SU(3) flavor symmetry into weak interaction hadronic current, that is, in SU(3) transformation when hadrons in the same octet transform to each other the corresponding hadronic current transform accordingly. So the hadronic current of $\Delta S = 0$

can transform into that of $\Delta S = 1$. For semi-leptonic decay, the Hamiltonian is

$$H_{\mathrm{W}} = \frac{G_{\mathrm{F}}}{\sqrt{2}}(J_\mu^+ j^{-\mu} + \mathrm{h.c.}) \tag{7.6.1}$$

Cabibbo thought, J_μ^+ (Positively charged hadronic current) should include both $\Delta S = 0$ and $\Delta S = 1$ currents, i.e.

$$J_\mu^+ = J_\mu^{+,\Delta S=0} \cos\theta_{\mathrm{C}} + J_\mu^{+,\Delta S=1} \sin\theta_{\mathrm{C}} \tag{7.6.2}$$

At quark level, Eq. (7.6.2) can be written as

$$J_\mu^+ = \bar{\mathrm{u}}\gamma_\mu(1-\gamma_5)\,(\mathrm{d}\cos\theta_{\mathrm{C}} + \mathrm{s}\sin\theta_{\mathrm{C}}) \tag{7.6.3}$$

Equation (7.6.1) read now

$$H_{\mathrm{W}} = \frac{G_{\mathrm{F}}}{\sqrt{2}}\bar{\mathrm{u}}\gamma_\mu(1-\gamma_5)\,(\mathrm{d}\cos\theta_{\mathrm{C}} + \mathrm{s}\sin\theta_{\mathrm{C}})\,\bar{l}\gamma^\mu\,(1-\gamma_5)\,\nu_l + \mathrm{h.c.} \tag{7.6.4}$$

From Eq. (7.6.4), we can see that weak interaction eigenstates are not d or s quarks, but are their linear combination

$$\mathrm{d}' = \mathrm{d}\cos\theta_{\mathrm{C}} + \mathrm{s}\sin\theta_{\mathrm{C}} \tag{7.6.5}$$

If take s′ as another weak interaction eigenstate orthogonal to d′, then

$$\mathrm{s}' = -\mathrm{d}\sin\theta_{\mathrm{C}} + \mathrm{s}\cos\theta_{\mathrm{C}} \tag{7.6.6}$$

So the mixing of $(\mathrm{d},\mathrm{s}) \to (\mathrm{d}',\mathrm{s}')$ can be written as

$$\begin{pmatrix} \mathrm{d}' \\ \mathrm{s}' \end{pmatrix} = \begin{pmatrix} \cos\theta_{\mathrm{C}} & \sin\theta_{\mathrm{C}} \\ -\sin\theta_{\mathrm{C}} & \cos\theta_{\mathrm{C}} \end{pmatrix} \begin{pmatrix} \mathrm{d} \\ \mathrm{s} \end{pmatrix} \tag{7.6.7}$$

where u, d, s are quark mass eigenstates, while d′, s′ are weak interaction eigenstates, these two eigenstates do not coincide. This concept of quark mixing is proposed first by Cabibbo in history.

Equation (7.6.2) explains why in Eqs. (7.3.3), (7.5.14), (7.5.21) and (7.5.22) appears $\cos\theta_{\mathrm{C}}$, while in Eq. (7.5.27) appears $\sin\theta_{\mathrm{C}}$.

From Eq. (7.6.4) we see that only d′ participates weak interactions, while s′ in Eq. (7.6.6) does not. This is because that in Eq. (7.6.1) and Eq. (7.6.4) only the weak interaction Hamiltonian of the charged current is presented, while the Hamiltonian of the neutral current weak interactions is not given. In fact, in addition to the charged current of Eq. (7.6.3), there should be neutral current term like

$$J_\mu^0 = \bar{\mathrm{u}}\gamma_\mu(1-\gamma_5)\mathrm{u} + \bar{\mathrm{d}}\gamma_\mu(1-\gamma_5)\mathrm{d} \tag{7.6.8}$$

If only d′ participates weak interactions, the neutral current term is

$$J_\mu^0 = \bar{\mathrm{u}}\mathrm{u} + \bar{\mathrm{d}}'\mathrm{d}' = \bar{\mathrm{u}}\mathrm{u} + \left(\cos^2\theta_\mathrm{C}\bar{\mathrm{d}}\mathrm{d} + \sin^2\theta_\mathrm{C}\bar{\mathrm{s}}\mathrm{s}\right) + \sin\theta_\mathrm{C}\cos\theta_\mathrm{C}\left(\bar{\mathrm{d}}\mathrm{s} + \bar{\mathrm{s}}\mathrm{d}\right) \tag{7.6.9}$$

where for convenience we neglect the factor of $\gamma_\mu(1-\gamma_5)$. The last term in Eq. (7.6.9) involves flavor changing neutral current $\bar{\mathrm{d}}\mathrm{s}$, $\bar{\mathrm{s}}\mathrm{d}$, these terms will induce weak decays with the same order of magnitude like that induced by charged current. For example, the decay

$$\mathrm{K}^+ \to \pi^+ + \bar{l} + l, \quad l = \mathrm{e}, \mu \tag{7.6.10}$$

will have quite large branching fraction. But in experiments the branching fraction of the decays of Eq. (7.6.10) is less than 10^{-7}.

In order to solve this difficulty caused by flavor-changing neutral current, in 1970 Glashow, Iliopoulos and Maiani [8] proposed a theory including the fourth quark, so this difficulty was solved.

Before 1970, we knew already that there were two lepton doublets participating weak interactions, i.e.

$$\begin{pmatrix}\nu_\mathrm{e}\\ \mathrm{e}\end{pmatrix}, \quad \begin{pmatrix}\nu_\mu\\ \mu\end{pmatrix} \tag{7.6.11}$$

But the quark doublets participating weak interactions were

$$\begin{pmatrix}\mathrm{u}\\ \mathrm{d}'\end{pmatrix}, \quad \begin{pmatrix}?\\ \mathrm{s}'\end{pmatrix} \tag{7.6.12}$$

On the top of the s′ something seems missing. So the three of GIM proposed the existence of the fourth flavored quark c based on symmetry consideration. Its charge is $\frac{2}{3}$ of the proton charge like the u quark. Thus Eq. (7.6.12) can be rewritten as

$$\begin{pmatrix}\mathrm{u}\\ \mathrm{d}'\end{pmatrix}, \quad \begin{pmatrix}\mathrm{c}\\ \mathrm{s}'\end{pmatrix} \tag{7.6.13}$$

Equation (7.6.9) extends to

$$\begin{aligned} J_\mu^0 &= \bar{\mathrm{u}}\mathrm{u} + \bar{\mathrm{c}}\mathrm{c} + \bar{\mathrm{d}}'\mathrm{d}' + \bar{\mathrm{s}}'\mathrm{s}' \\ &= \bar{\mathrm{u}}\mathrm{u} + \bar{\mathrm{c}}\mathrm{c} + \bar{\mathrm{d}}\mathrm{d} + \bar{\mathrm{s}}\mathrm{s} \end{aligned} \tag{7.6.14}$$

In that way in Eq. (7.6.14) the flavor changing neutral currents no longer appear. This method for canceling flavor changing neutral currents is called GIM mechanism. So before the discovery of the fourth quark c, it was already proposed by GIM (c quark was discovered in 1974). GIM mechanism cancels flavor changing neutral currents just at the lowest order of weak interactions (at the order of G_{F}), in high order of weak interactions flavor changing neutral currents can still possibly appear, so GIM mechanism is also called GIM suppression.

After introducing c quark, the charged current can be written as

$$J_\mu^+ = (\bar{\mathrm{u}}, \bar{\mathrm{c}})\,\gamma_\mu\,(1-\gamma_5)\,V_{\mathrm{C}}\begin{pmatrix}\mathrm{d}\\ \mathrm{s}\end{pmatrix} \tag{7.6.15}$$

where

$$V_{\mathrm{C}} = \begin{pmatrix}\cos\theta_{\mathrm{C}} & \sin\theta_{\mathrm{C}}\\ -\sin\theta_{\mathrm{C}} & \cos\theta_{\mathrm{C}}\end{pmatrix} \tag{7.6.16}$$

is called Cabibbo quark mixing matrix.

In 1974 S.C.C. Ting group and B. Richter group [19] discovered J/ψ particle at the Brookhaven National Laboratory (BNL) and the Stanford Linear Accelerator Center(SLAC), respectively. After about one year discussion by theorists and experimentalists they finally identified this new particle as a bound state of $\mathrm{c}\bar{\mathrm{c}}$, therefore the existence of the c quark is confirmed.

7.7. Kobayashi–Maskawa Model

In 1973 Kobayashi and Maskawa [9] extended Cabibbo's idea of two generation quark mixing and of the GIM mechanism to the case of three generation quarks, i.e. in addition to (u, d) and (c, s) two generation quarks they introduced the third generation quarks (t, b), t quark charge is $\frac{2}{3}$, b charge is $\frac{-1}{3}$. The three generation quarks participating weak interactions are

$$\begin{pmatrix}\mathrm{u}\\ \mathrm{d}'\end{pmatrix},\quad \begin{pmatrix}\mathrm{c}\\ \mathrm{s}'\end{pmatrix},\quad \begin{pmatrix}\mathrm{t}\\ \mathrm{b}'\end{pmatrix} \tag{7.7.1}$$

where

$$\begin{pmatrix}\mathrm{d}'\\ \mathrm{s}'\\ \mathrm{b}'\end{pmatrix} = \begin{pmatrix}V_{\mathrm{ud}} & V_{\mathrm{us}} & V_{\mathrm{ub}}\\ V_{\mathrm{cd}} & V_{\mathrm{cs}} & V_{\mathrm{cb}}\\ V_{\mathrm{td}} & V_{\mathrm{ts}} & V_{\mathrm{tb}}\end{pmatrix}\begin{pmatrix}\mathrm{d}\\ \mathrm{s}\\ \mathrm{b}\end{pmatrix} \tag{7.7.2}$$

Here d, s, b are quark mass eigenstates and d′, s′, b′ are weak interaction eigenstates which are mixing states of the mass eigenstates. The mixing matrix

$$V_{\mathrm{CKM}} \equiv \begin{pmatrix} V_{\mathrm{ud}} & V_{\mathrm{us}} & V_{\mathrm{ub}} \\ V_{\mathrm{cd}} & V_{\mathrm{cs}} & V_{\mathrm{cb}} \\ V_{\mathrm{td}} & V_{\mathrm{ts}} & V_{\mathrm{tb}} \end{pmatrix} \tag{7.7.3}$$

in called Cabibbo–Kobayashi–Maskawa matrix, simply called CKM matrix. The difference of Eq. (7.7.3) and Cabibbo mixing matrix V_{C} (see Eq. (7.6.16)) is that it changes from 2 × 2 matrix into 3 × 3 one. V_{CKM} is 3 × 3 complex unitary matrix, when it appears in weak interaction charged current,

$$J_\mu^+ = (\bar{\mathrm{u}}, \bar{\mathrm{c}}, \bar{\mathrm{t}})\, \gamma_\mu \,(1-\gamma_5)\, V_{\mathrm{CKM}} \begin{pmatrix} \mathrm{d} \\ \mathrm{s} \\ \mathrm{b} \end{pmatrix} \tag{7.7.4}$$

V_{CKM} should have nine independent real parameters. From Eq. (7.7.4), we see that the six fields $(\bar{\mathrm{u}}, \bar{\mathrm{c}}, \bar{\mathrm{t}})$ and $(\mathrm{d}, \mathrm{s}, \mathrm{b})$ have five independent phases which can be chosen arbitrarily, so that five real parameters of V_{CKM} can be canceled, thus in V_{CKM} only four independent real parameters are left. We can choose three of them as rotation angles θ_{12}, θ_{23}, θ_{31}, the other as a complex phase δ_{13}. The existence of this complex phase makes $V_{\mathrm{CKM}}^* \neq V_{\mathrm{CKM}}$, this will induce CP violation. This is extremely important result. For the parameterization of V_{CKM}, CP violation can be found in the book of CP Non-conservation [20]. We first give two parameterization schemes used in general, i.e. the standard scheme [13] and the Wolfenstein scheme [21]. Because every generation quarks appears in pair and V_{CKM} is unitary, so there are no flavor changing neutral currents to happen.

The standard scheme of the V_{CKM} parameterization can be expressed as a product of three rotation matrices

$$\begin{aligned} V_{\mathrm{CKM}} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}\mathrm{e}^{-\mathrm{i}\delta_{13}} \\ 0 & 1 & 0 \\ -s_{13}\mathrm{e}^{\mathrm{i}\delta_{13}} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}\mathrm{e}^{-\mathrm{i}\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}\mathrm{e}^{\mathrm{i}\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}\mathrm{e}^{\mathrm{i}\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}\mathrm{e}^{\mathrm{i}\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}\mathrm{e}^{\mathrm{i}\delta_{13}} & c_{23}c_{13} \end{pmatrix} \end{aligned} \tag{7.7.5}$$

where $c_{ij} \equiv \cos\theta_{ij}$, $s_{ij} \equiv \sin\theta_{ij}$, i, $j = 1, 2, 3$, and δ_{13} is a phase.

The Wolfenstein parametrization is as follows:

$$V_{\mathrm{CKM}} \approx \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho-\mathrm{i}\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-\mathrm{i}\eta) & -A\lambda^2 & 1 \end{pmatrix} + O(\lambda^4) \quad (7.7.6)$$

At present, the experiments give that

$$\begin{aligned} &\lambda = 0.22535 \pm 0.00065, && A = 0.811^{+0.022}_{-0.012} \\ &\bar{\rho} = 0.131^{+0.026}_{-0.013}, && \bar{\eta} = 0.345^{+0.013}_{-0.014} \\ &\bar{\rho} = \rho(1-\lambda^2/2), && \bar{\eta} = \eta(1-\lambda^2/2) \end{aligned} \quad (7.7.7)$$

7.8. The Limitation of the Four Fermion Point-like Interactions and the Intermediate Vector Bosons

In the theory of the four fermion current–current coupling, the four fermions interact at one point of space–time. Although this theory get many greet achievements in describing weak interaction experiments, but it has yet limitations. For example, for the following inelastic scattering process

$$\nu_\mu + \mathrm{e}^- \to \mu^- + \nu_\mathrm{e} \quad (7.8.1)$$

when we consider only s wave, the calculated cross section is

$$\sigma\left(\nu_\mu + \mathrm{e}^- \to \mu^- + \nu_\mathrm{e}\right) = \frac{8}{\pi}\left(G_\mathrm{F} k\right)^2 \quad (7.8.2)$$

where $k = \left|\overrightarrow{k}\right|$ is the ν_μ momentum in the center-of-mass system. When $k \to \infty$ the scattering cross section in Eq. (7.8.2) diverges. According to unitary bound of s wave scattering (probability conservation), we should have

$$\sigma \leq \frac{\pi}{k^2} \quad (7.8.3)$$

So the probability conservation demands

$$\frac{8}{\pi}\left(G_\mathrm{F} k\right)^2 \leq \frac{\pi}{k^2} \quad (7.8.4)$$

This will lead to

$$k \leq \left(\frac{\pi}{2\sqrt{2}G_\mathrm{F}}\right)^{\frac{1}{2}} \sim 300\,\mathrm{GeV} \quad (7.8.5)$$

Equation (7.8.5) indicates that the four fermion current–current coupling theory of weak interactions will break the probability conservation law when the momentum $k \geq 300\,\text{GeV}$, so is not applicable. So, the theory of weak interactions must be revised at high energy. Comparing with electromagnetic interactions which are realized through interchanging photons, weak interactions can be realized through interchanging massive vector particles. T.D. Lee, M. Rosenbluth and C.N. Yang [11] proposed a intermediate vector boson theory of weak interactions according to this idea. They assumed that the charged weak currents interact through interchanging intermediate vector bosons $\mathrm{W}^{\pm}$, thus the effective Hamiltonian of the weak interactions becomes

$$H_{\mathrm{W}} = \frac{g}{2\sqrt{2}}\left(J_\mu^+ W^{+\mu} + j_\mu^+ W^{+\mu}\right) + \text{h.c.} \tag{7.8.6}$$

Note that in the expansion of the $W^{+\mu}$ field the field operators annihilate W^+ and create W^- particles. So Eq. (7.8.6) describes a neutral operator and

$$J_\mu^+ = (\bar{\mathrm{u}}, \bar{\mathrm{c}}, \bar{\mathrm{t}})\,\gamma_\mu\,(1-\gamma_5)\,V_{\mathrm{CKM}}\begin{pmatrix}\mathrm{d}\\ \mathrm{s}\\ \mathrm{b}\end{pmatrix} \tag{7.8.7}$$

$$j_\mu^+ = (\bar{\nu}_{\mathrm{e}}, \bar{\nu}_\mu, \bar{\nu}_\tau)\,\gamma_\mu\,(1-\gamma_5)\begin{pmatrix}\mathrm{e}\\ \mu\\ \tau\end{pmatrix} \tag{7.8.8}$$

Consider the neutrino-quark scattering process

$$\nu + \mathrm{d} \to \mathrm{u} + l \tag{7.8.9}$$

According to the theory of the intermediate vector bosons, the corresponding Feynman diagram is shown in Fig. 7.4.

Using Eqs. (7.8.6)–(7.8.8), the amplitude of the scattering process in Fig. 7.4 can be written as

$$M \propto \frac{g^2}{(2\sqrt{2})^2} V_{\mathrm{ud}}\bar{\mathrm{u}}\gamma_\mu(1-\gamma_5)d\frac{-\mathrm{i}}{p^2 - m_{\mathrm{W}}^2}\bar{l}\gamma^\mu(1-\gamma_5)\nu \tag{7.8.10}$$

Obviously, when the exchanging momentum is small, p^2 can be neglected. We come back to the low-energy four fermion interaction, so we have

$$\frac{G_{\mathrm{F}}}{\sqrt{2}} = \frac{g^2}{8m_{\mathrm{W}}^2} \tag{7.8.11}$$

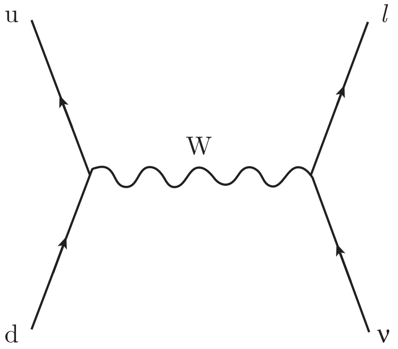


Fig. 7.4. The Feynman diagram for $\nu + \mathrm{d} \to \mathrm{u} + l$ scattering.

This is compatible with the prediction of the unified theory of the electroweak interactions in Chapter 8.

Although the assumption of the intermediate vector boson can solve the problem of breaking probability conservation for the point-like four fermion interactions at high energy, but the new theory itself cannot be renormalized so must be corrected. Only after introducing gauge invariance and spontaneous symmetry breaking then the above difficulty can be solved. That is the topic of Chapter 8.

7.9. Conservation of Vector Current

The assumption of the conservation of vector current(CVC) was proposed in the late 1950s of the 20th century [6, 22]. If we write down the effective Hamiltonian for charged semi-leptonic weak decay

$$H_{\mathrm{cc}}^{\mathrm{SL}} = \frac{G_{\mathrm{F}}}{\sqrt{2}}[J_\mu^+ j^{-\mu} + J_\mu^- j^{+\mu}] \tag{7.9.1}$$

where the hadronic current $J_\mu^\pm = V_\mu^\pm - A_\mu^\pm$ is the difference of the vector current $V_\mu^\pm$ and the axial vector current $A_\mu^\pm$ V–A theory, $j_\mu^\pm$ is the leptonic current. CVC assumption includes two points:

(1) $$\frac{V_\mu^+}{\sqrt{2}\cos\theta_{\mathrm{C}}}, \quad (J_\mu^{\mathrm{em}})_{I=1}, \quad \frac{V_\mu^-}{\sqrt{2}\cos\theta_{\mathrm{C}}}$$

form isospin triplet, $I_3 = 1, 0, -1$, where $(J_\mu^{\mathrm{em}})_{I=1}$ is the isospin vector part of the hadronic electromagnetic current.

(2) The above three current components all conserve, i.e.

$$\frac{\partial V_\mu^\pm}{\partial x_\mu} = 0, \quad \frac{\partial}{\partial x_\mu}(J_\mu^{\text{em}})_{I=1} = 0$$

The assumption (1) is understandable, because we already know today that $V_\mu^\pm$ and the hadronic electromagnetic current J_μ^{em} are all vector currents consisted of quarks and antiquarks. The important physical meaning is that the weak interactions and the electromagnetic interactions have internal connection at hadronic level. This is actually one of the foundations of the unified theory of the electroweak interactions. The assumption (2) is obviously approximate from present point of view. In fact, using quark language

$$V_\mu^+ = \bar{\text{u}}\gamma_\mu \text{d}, \quad V_\mu^- = \bar{\text{d}}\gamma_\mu \text{u}, \quad (J_\mu^{\text{em}})_{I=1} = \frac{1}{\sqrt{2}}(\bar{\text{d}}\gamma_\mu \text{d} - \bar{\text{u}}\gamma_\mu \text{u}) \tag{7.9.2}$$

where the form of $(J_\mu^{\text{em}})_{I=1}$ can be understood from that of π^0 in quark model, $\pi^0 = \frac{\bar{\text{d}}\text{d} - \bar{\text{u}}\text{u}}{\sqrt{2}}$ is just $I = 1$, $I_3 = 0$ state, so the last term of Eq. (7.9.2) is just the electromagnetic current $(J_\mu^{\text{em}})_{I=1,I_3=0}$.

$$\partial_\mu V^{+\mu} = \partial_\mu(\bar{\text{u}}\gamma_\mu \text{d}) = (\partial_\mu \bar{\text{u}})\,\gamma_\mu \text{d} + \bar{\text{u}}\gamma^\mu \partial_\mu \text{d} \tag{7.9.3}$$

According to Dirac equation when there are strong interactions

$$(\text{i}\, \not{D} - m)\,\psi(x) = 0 \tag{7.9.4}$$

where

$$\not{D} = D_\mu \gamma^\mu, \quad D_\mu = \partial_\mu - \text{i}g A_\mu^a \frac{\lambda_a}{2} \tag{7.9.5}$$

λ_a is Gell-Mann SU(3)matrix. From Eqs. (7.9.4) and (7.9.5), we obtain

$$\begin{aligned} \left(\text{i}\partial_\mu \gamma^\mu + g A_\mu^a \frac{\lambda_a}{2}\gamma^\mu - m\right)\psi &= 0 \\ \bar{\psi}\left(-\text{i}\overleftarrow{\partial}_\mu \gamma^\mu + g A_\mu^a \frac{\lambda_a}{2}\gamma^\mu - m\right) &= 0 \end{aligned} \tag{7.9.6}$$

In Eq. (7.9.6) take $\psi = \text{u}$ or d, we can solve out $\partial_\mu \bar{\text{u}}\gamma_\mu$ and $\gamma^\mu \partial_\mu \text{d}$. Substituting them into Eq. (7.9.3) we obtain

$$\partial_\mu V^{+\mu} = (m_{\text{d}} - m_{\text{u}})\,\bar{\text{u}}\text{d} \tag{7.9.7}$$

where m_{d}, m_{u} are the quark current masses of d, u, respectively. Equation (7.9.7) shows that $\partial_\mu V^{\pm\mu} \neq 0$. But the quark current masses are very small (order of MeV), so we have

$$\partial_\mu V^{\pm\mu} \approx 0 \tag{7.9.8}$$

Similarly, it is easy to prove that the conservation of the hadronic electromagnetic current is accurate. This can be seen from Eq. (7.9.7), because in the right-hand side of the partial derivative of $(J_\mu^{\text{em}})_{I=1}$, the two subtracted quark masses are equal, so is accurately zero, i.e.

$$\partial_\mu (J_\mu^{\text{em}})_{I=1} \equiv 0 \tag{7.9.9}$$

Note that for the case of including the strong, weak, and electromagnetic interactions Eq. (7.9.7) still holds. The only change in Eq. (7.9.6) is the covariant derivative D_μ, but the result of Eq. (7.9.7) does not change.

The CVC assumption has been tested early in the nucleus β decay, the corresponding example of experiments is given in T.D. Lee's book [23]. Now we give the example of the π meson β decay for testing CVC assumption (1).

The π^- meson β decay is given by

$$\pi^- \to \pi^0 + \mathrm{e}^- + \bar{\nu}_\mathrm{e} \tag{7.9.10}$$

Its hadronic matrix element $\langle \pi^0(k') | V_\mu^+ | \pi^-(k) \rangle$ can connect with the electromagnetic current J_μ^{em} through the CVC assumption (1), i.e.

$$\begin{aligned} \left\langle \pi^0 (k') \right| V_\mu^+ \left| \pi^- (k) \right\rangle &= -\sqrt{2}\cos\theta_\mathrm{C} \left\langle \pi^- (k') \right| J_\mu^{\text{em}} \left| \pi^- (k) \right\rangle \\ &= -\cos\theta_\mathrm{C}\sqrt{2}F_\pi\left(q^2\right)(k+k')_\mu \end{aligned} \tag{7.9.11}$$

where $F_\pi\left(q^2\right)$ is the π meson electromagnetic form factor, $q = k' - k$. Because $\Delta = m_{\pi^\pm} - m_{\pi^0} \sim 4\text{MeV}$ which is very small, so q^2 is also very small, $F_\pi\left(q^2\right) \approx F_\pi(0) = 1$. From Eq. (7.9.11) we can compute the decay width of the decay of Eq. (7.9.10) as

$$\Gamma = G_\mathrm{F}^2 \frac{\cos^2\theta_\mathrm{C}}{30\pi^3}\left[1 - \frac{\Delta}{2m_\pi}\right]\Delta^5 \bar{F}(1+\delta_\pi) \tag{7.9.12}$$

where $\bar{F}$ is the phase space factor, $\bar{F} \approx 1$, The radiation correction factor is given by $\delta_\pi \sim 1\%$ [24]. So we can compute out the branching fraction of the decay in Eq. (7.9.10) [25]:

$$\text{Br}\left(\pi^- \to \pi^0 \mathrm{e}^- \bar{\nu}_\mathrm{e}\right) = (1.0482 \pm 0.0048) \times 10^{-8}$$

While the measured branching fraction is given by $(1.026 \pm 0.039) \times 10^{-8}$ which is compatible with the theoretical result [26].

There are other more recent measurements and applications. The readers can refer the recent review [26] and references therein.

7.10. Chiral Symmetry Breaking and PCAC

Given quark field $q_i^\alpha(x)$, $i = 1, 2, 3$, represents "flavor" indices u, d, s, $\alpha = 1, 2, 3$, denote "color" indices, then $q_i^\alpha(x)$ can be written as a column matrix

$$q(x) = \begin{pmatrix} \mathrm{u}^\alpha(x) \\ \mathrm{d}^\alpha(x) \\ \mathrm{s}^\alpha(x) \end{pmatrix} \tag{7.10.1}$$

From the above expressions, we can write down the vector currents and the axial vector currents

$$\begin{aligned} V_\mu^a(x) &= \bar{q}(x)\gamma_\mu \frac{\lambda^a}{2} q(x) \\ A_\mu^a(x) &= \bar{q}(x)\gamma_\mu\gamma_5 \frac{\lambda^a}{2} q(x) \end{aligned} \tag{7.10.2}$$

where $\lambda^a/2$, $a = 1, 2, \ldots, 8$ are SU(3) group generators, the color index α is not given manifestly. Put the quark mass matrix as

$$M = \mathrm{diag}(m_\mathrm{u}, m_\mathrm{d}, m_\mathrm{s}) \tag{7.10.3}$$

Using QCD equation of motion (7.9.6), and considering that $A_\mu^a \lambda^a/2$ are color space SU(3) group operators which have nothing to do with flavor space SU(3) group generators λ^a in Eq. (7.10.2), we have

$$\partial^\mu V_\mu^a(x) = \mathrm{i}\bar{q}(x)\gamma_\mu \left(M\frac{\lambda^a}{2} - \frac{\lambda^a}{2}M \right) q(x) \tag{7.10.4}$$

$$\partial^\mu A_\mu^a(x) = \mathrm{i}\bar{q}(x) \left(M\frac{\lambda^a}{2} + \frac{\lambda^a}{2}M \right) q(x) \tag{7.10.5}$$

Assume that u, d, s quark masses are all zero, from the above two equations we see that both the vector current and the axial vector current conserve, i.e.

$$\begin{aligned} \partial^\mu V_\mu^a(x) &= 0 \\ \partial^\mu A_\mu^a(x) &= 0 \end{aligned} \tag{7.10.6}$$

So we get the conserved charges $Q_V^a(t)$, $Q_A^a(t)$

$$\begin{aligned} Q_V^a(t) &= \int \mathrm{d}^3x V_0^a(\vec{x},t), \quad \frac{\partial}{\partial t}Q_V^a(t) = 0 \\ Q_A^a(t) &= \int \mathrm{d}^3x A_0^a(\vec{x},t), \quad \frac{\partial}{\partial t}Q_A^a(t) = 0 \end{aligned} \tag{7.10.7}$$

Using the following equal time commutators

$$\begin{aligned} &\left[\int \mathrm{d}^3x'\psi^+(\vec{x}',t)\Gamma'\psi(\vec{x}',t), \int \mathrm{d}^3x\psi^+(\vec{x},t)\Gamma\psi(\vec{x},t)\right] \\ &\quad = \int \mathrm{d}^3x\psi^+(\vec{x},t)[\Gamma',\Gamma]\psi(\vec{x},t) \end{aligned} \tag{7.10.8}$$

we can derive the following equal time ($t = t'$) commutation relation

$$\begin{aligned} [V_0^a(x), V_\mu^b(x')] &= [A_0^a(x), A_\mu^b(x')] = -f^{abc}V_\mu^c(x)\delta^3(\vec{x}-\vec{x}') \\ [V_0^a(x), A_\mu^b(x')] &= [A_0^a(x), V_\mu^b(x')] = -f^{abc}A_\mu^c(x)\delta^3(\vec{x}-\vec{x}') \end{aligned} \tag{7.10.9}$$

where $x = (\vec{x},t)$, $x' = (\vec{x}',t')$, f^{abc} are SU(3) group structure constants (see Table 5.1).

Using Eqs. (7.10.7) and (7.10.9), it is easy to derive that

$$\begin{aligned} [Q_V^a(t), V_\mu^b(x)] &= [Q_A^a(t), A_\mu^b(x)] = \mathrm{i}f^{abc}V_\mu^c(x) \\ [Q_V^a(t), A_\mu^b(x)] &= [Q_A^a(t), V_\mu^b(x)] = \mathrm{i}f^{abc}A_\mu^c(x) \end{aligned} \tag{7.10.10}$$

In Eq. (7.10.10) take $\mu = 0$ and integrate $\int \mathrm{d}^3x$ we have

$$\begin{aligned} [Q_V^a(t), Q_V^b(t)] &= [Q_A^a(t), Q_A^b(t)] = \mathrm{i}f^{abc}Q_V^c(t) \\ [Q_V^a(t), Q_A^b(t)] &= [Q_A^a(t), Q_V^b(t)] = \mathrm{i}f^{abc}Q_A^c(t) \end{aligned} \tag{7.10.11}$$

Define

$$Q_\pm^a = \frac{1}{2}(Q_V^a \pm Q_A^a) \tag{7.10.12}$$

then we obtain

$$\begin{aligned} [Q_+^a(t), Q_+^b(t)] &= \mathrm{i}f^{abc}Q_+^c(t) \\ [Q_-^a(t), Q_-^b(t)] &= \mathrm{i}f^{abc}Q_-^c(t) \\ [Q_+^a(t), Q_-^b(t)] &= 0 \end{aligned} \tag{7.10.13}$$

From the above equations we see that $Q_+^a(t)$, $Q_-^a(t)$ ($a = 1,2,\ldots,8$) form SU(3) group generators, respectively. The conserved charges $Q_\pm^a(t)$

certainly commute with the Hamiltonian H_{QCD}. So the system possesses SU(3) × SU(3) chiral symmetry. The reason for calling it chiral symmetry is that this SU(3) × SU(3) symmetry holds only when the three quark masses m_{u},m_{d},m_{s} are equal to zero. In that case, the left-handed components $q_{\mathrm{L}}(x)$ and the right-handed components $q_{\mathrm{R}}(x)$ fulfill SU(3) symmetry independently, i.e. the system possesses $\mathrm{SU(3)_L} \times \mathrm{SU(3)_R}$ symmetry.

$$J^a_{\mu\mathrm{L}} = \bar{q}_{\mathrm{L}}(x)\gamma_\mu \frac{\lambda_a}{2} q_{\mathrm{L}}(x), \quad J^a_{\mu\mathrm{R}} = \bar{q}_{\mathrm{R}}(x)\gamma_\mu \frac{\lambda_a}{2} q_{\mathrm{R}}(x) \tag{7.10.14}$$

are all conserved currents, i.e.

$$\partial^\mu J^a_{\mu\mathrm{L}}(x) = 0, \quad \partial^\mu J^a_{\mu\mathrm{R}}(x) = 0 \tag{7.10.15}$$

The corresponding conserved charges $Q^a_{\mathrm{L}}(t)$ and $Q^a_{\mathrm{R}}(t)$ are just $Q^a_-(t)$ and $Q^a_+(t)$ mentioned before, i.e.

$$Q^a_{\mathrm{L}}(t) = Q^a_-(t), \quad Q^a_{\mathrm{R}}(t) = Q^a_+(t) \tag{7.10.16}$$

The previous SU(3) × SU(3) symmetry and $\mathrm{SU(3)_L} \times \mathrm{SU(3)_R}$ symmetry here are the same symmetry.

Using the property of the spinor field under space inversion, it is easy to prove that

$$\begin{aligned} \hat{P}Q^a_+(t)\hat{P}^{-1} &= Q^a_-(t) \\ \hat{P}Q^a_-(t)\hat{P}^{-1} &= Q^a_+(t) \end{aligned} \tag{7.10.17}$$

So the parity transformation connect these two SU(3) symmetries.

We need to point out that the commutation relations of Eqs. (7.10.11)–(7.10.13) are derived from Eq. (7.10.8). But Eq. (7.10.8) is not rigorous, because when $t = t'$ it involves the product of four fermion fields at the same point of space time, so it has singularity. But we can prove that [see 23, Chapter 24, p. 774]) the only modification for Eqs. (7.10.9) and (7.10.10) is

$$\begin{aligned} &[V^a_0(\vec{x},t), V^b_i(\vec{x'},t)] = [A^a_0(\vec{x},t), A^b_i(\vec{x'},t)] \\ &\quad = -f^{abc}V^c_i(\vec{x},t)\delta^3(\vec{x}-\vec{x}') + \lambda\delta^{ab}\nabla_i\delta^3(\vec{x}-\vec{x}') \end{aligned} \tag{7.10.18}$$

where $i = 1,2,3$, $\nabla_i = \frac{\partial}{\partial x^i}$, the other commutation relations will not be changed. The term of $\lambda\delta^{ab}\nabla_i\delta^3(\vec{x}-\vec{x}')$ is called Schwinger term. Obviously, the existence of the Schwinger term does not affect the commutation relations after integration with respect to the space coordinates, so Eqs. (7.10.10)–(7.10.13) are not changed.

From the above discussions, we know that if we put quark masses to zero, then the vector and axial vector currents in Eq. (7.10.2) all conserve, i.e. both CVC and CAC all hold. But actually quark masses are not zero.

We first discuss the case where only u, d quarks exist. When $m_{\mathrm{u}} = m_{\mathrm{d}} = 0$ we only have SU(2) × SU(2) chiral symmetry. But when $m_{\mathrm{u}} = m_{\mathrm{d}} \neq 0$, Eqs. (7.10.4)–(7.10.5) degenerate into

$$V_\mu^i(x) = \bar{q}(x)\gamma_\mu \frac{\tau^i}{2} q(x), \quad A_\mu^i(x) = \bar{q}(x)\gamma_\mu\gamma_5 \frac{\tau^i}{2} q(x) \tag{7.10.19}$$

$$\partial^\mu V_\mu^i(x) = 0, \quad \partial^\mu A_\mu^i(x) = \mathrm{i}2m_{\mathrm{u}}\bar{q}(x)\gamma_5 \frac{\tau^i}{2} q(x) \tag{7.10.20}$$

where τ^i $(i = 1, 2, 3)$ are Pauli matrices, and

$$q(x) = \begin{pmatrix} \mathrm{u}^a(x) \\ \mathrm{d}^a(x) \end{pmatrix} \tag{7.10.21}$$

Because $\partial^\mu A_\mu^i(x) \neq 0$, so the corresponding charge $Q_A^i(t)$ no longer conserve, thus the SU(2) × SU(2) chiral symmetry is broken, only the SU(2) isospin symmetry is left, this is because $m_{\mathrm{u}} = m_{\mathrm{d}}$ still holds. Of course, the vector currents still conserve.

Put

$$\begin{aligned} \left|\pi^+\right\rangle &= \frac{1}{\sqrt{2}}\left(\left|\pi^1\right\rangle + \mathrm{i}\left|\pi^2\right\rangle\right) \\ \left|\pi^-\right\rangle &= \frac{1}{\sqrt{2}}\left(\left|\pi^1\right\rangle - \mathrm{i}\left|\pi^2\right\rangle\right) \\ \left|\pi^0\right\rangle &= \left|\pi^3\right\rangle \end{aligned} \tag{7.10.22}$$

where $\left|\pi^i\right\rangle$ $(i = 1, 2, 3)$ are vectors of the SU(2) isospin space.

Put the axial currents as

$$\begin{aligned} A_\mu^+ &= A_\mu^1 + \mathrm{i}A_\mu^2 \\ A_\mu^- &= A_\mu^1 - \mathrm{i}A_\mu^2 \end{aligned} \tag{7.10.23}$$

where the axial current $A_\mu^+(x)$ is positively charged current, i.e. it annihilates negatively charged particles and creates positively charged particles. Using Eq. (7.5.20), we get

$$\langle 0| A_\mu^\pm(x) \left|\pi^\mp\right\rangle = \mathrm{i}F_\pi k_\mu \mathrm{e}^{-\mathrm{i}k\cdot x} \tag{7.10.24}$$

so we can obtain

$$\langle 0| \partial^\mu A_\mu^\pm(x) \left|\pi^\mp\right\rangle = F_\pi m_\pi^2 \mathrm{e}^{-\mathrm{i}k\cdot x} \tag{7.10.25}$$

In Eq. (7.10.25), we have used π meson momentum squared $k_\mu k^\mu = m_\pi^2$.

$$\langle 0|\, \partial^\mu A_\mu^i(x) \left|\pi^j\right\rangle = \delta^{ij} f_\pi m_\pi^2 \mathrm{e}^{-\mathrm{i}k\cdot x} \tag{7.10.26}$$

where $f_\pi = \frac{1}{\sqrt{2}} F_\pi \approx 93\,\mathrm{MeV}$ is the π meson decay constant.

We can now introduce phenomenologically a π meson interpolating field $\vec{\phi}_\pi(x)$ which satisfies

$$\langle 0|\, \phi^i(x) \left|\pi^j\right\rangle = \delta^{ij} \mathrm{e}^{-\mathrm{i}k\cdot x} \tag{7.10.27}$$

then we have

$$\partial^\mu \vec{A}_\mu(x) = f_\pi m_\pi^2 \vec{\phi}_\pi(x) \tag{7.10.28}$$

Because the π meson mass is very small, the right-hand side of Eq. (7.10.28) is also very small, so it can be viewed as partial conservation of axial vector current. A_μ is called partially conserved axial current (PCAC) [27]. Comparing Eqs. (7.10.28) and (7.10.20) we see that

$$\vec{\phi}_\pi(x) = \mathrm{i}\frac{m_\mathrm{u}}{f_\pi m_\pi^2} \bar{q}(x)\gamma_5 \vec{\tau} q(x) \tag{7.10.29}$$

It is easy to extrapolate the above discussion to the case of SU(3) × SU(3), and deduce the PCAC of K meson. But because $m_\mathrm{K}^2 \gg m_\pi^2$, this extrapolation is not meaningful.

7.11. Mixing of Neutral Mesons and *CP* Violation

In this section we discuss briefly the mixing and *CP* violation of K^0–$\bar{\mathrm{K}}^0$, B^0–$\bar{\mathrm{B}}^0$ and D^0–$\bar{\mathrm{D}}^0$. If only strong and electromagnetic interactions exist, then the neutral meson and its antiparticle have same mass. But when the weak interactions come in, neutral meson P^0 and its antiparticle $\bar{\mathrm{P}}^0$ will decay. If they have same decay final states, i.e. P^0 and $\bar{\mathrm{P}}^0$ can decay into same final states (real or virtual), then P^0 and $\bar{\mathrm{P}}^0$ can transform into each other. We call this mutual transformation as mixing or oscillation.

7.11.1. *Quantum mechanical description*

The state vector of the neutral meson P^0–$\bar{\mathrm{P}}^0$ system is

$$|\Phi(t)\rangle = a(t) \left|\mathrm{P}^0\right\rangle + b(t) \left|\bar{\mathrm{P}}^0\right\rangle + \sum_{i=1}^{\infty} c_i(t) \left|n_i\right\rangle \tag{7.11.1}$$

where t is the proper time in P^0, $\bar{\mathrm{P}}^0$ rest frame and $|n_i\rangle$ are the possible decay final states of P^0, $\bar{\mathrm{P}}^0$. Now $|\Phi(t)\rangle$ satisfies the Schrodinger equation

$$\mathrm{i}\hbar\frac{\partial}{\partial t}|\Phi(t)\rangle = \mathcal{H}|\Phi(t)\rangle \tag{7.11.2}$$

This is an equation in infinitely dimensional Hilbert space, The Hamiltonian $\mathcal{H}$ is an infinite-dimensional Hermitian operator in Hilbert space. We do not have any method now to solve his equation rigorously, but we are only interest in the following cases [28]:

(1) The initial state is the linear combination of $\left|\mathrm{P}^0\right\rangle$, $\left|\bar{\mathrm{P}}^0\right\rangle$, i.e.

$$|\psi(0)\rangle = a(0)\left|\mathrm{P}^0\right\rangle + b(0)\left|\bar{\mathrm{P}}^0\right\rangle$$

(2) We only want to know the time variation of the coefficients $a(t)$, $b(t)$ on time.

(3) Only consider the timescale t which is much larger than the strong interaction timescale, so the Weisskopf–Wigner approximation [29] can be used, and outside of the states $\left|\mathrm{P}^0\right\rangle$ and $\left|\bar{\mathrm{P}}^0\right\rangle$ will not be considered.

Under the above three conditions, the situation will be much simplified. The state vector can be written as

$$|\psi(t)\rangle = a(t)\left|\mathrm{P}^0\right\rangle + b(t)\left|\bar{\mathrm{P}}^0\right\rangle \tag{7.11.3}$$

In the subspace spanned by P^0, $\bar{\mathrm{P}}^0$, it can be expressed by a column matrix

$$|\psi(t)\rangle = \begin{pmatrix} a(t) \\ b(t) \end{pmatrix} \tag{7.11.4}$$

The corresponding Schrodinger equation is

$$\mathrm{i}\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = H|\psi(t)\rangle \tag{7.11.5}$$

where H is a 2×2 matrix which is not Hermitian, because P^0, $\bar{\mathrm{P}}^0$ will decay and the particle number does not conserve. Usually we write H as

$$H = M - \frac{\mathrm{i}}{2}\Gamma \tag{7.11.6}$$

H has three parts $H = H_{\mathrm{str}} + H_\gamma + H_{\mathrm{w}}$, where H_{str} is strong interaction part, H_γ is electromagnetic interaction part, H_{w} is weak interaction part.

$M^+ = M, \Gamma^+ = \Gamma$, they are all Hermitian matrices. From perturbation theory, we have

$$\begin{aligned} M_{ij} &= m_{\mathrm{p}}\delta_{ij} + \langle i| H_{\mathrm{w}} |j\rangle + \sum_n \mathcal{P}\frac{\langle i| H_{\mathrm{w}} |n\rangle \langle n| H_{\mathrm{w}} |j\rangle}{m_{\mathrm{P}} - E_n} \\ \Gamma_{ij} &= 2\pi \sum_n \delta(m_{\mathrm{P}} - E_{\mathrm{n}}) \langle i| H_{\mathrm{w}} |n\rangle \langle n| H_{\mathrm{w}} |j\rangle \end{aligned} \tag{7.11.7}$$

where i, j take P^0 or $\bar{\mathrm{P}}^0$, and m_{P} is the mass of P^0, $\bar{\mathrm{P}}^0$, and $\mathcal{P}$ denotes principal value operation. The derivation of Eq. (7.11.7) can be found in [30].

In order to discuss the properties of M_{ij} and Γ_{ij} under CPT, CP, and T transformations, we first look at the property of anti-unitary operators, CPT and T transformations belong to this category.

Any anti-unitary operator A can always be written as a product of a unitary operator U and a operation K for taking complex conjugation:

$$A = UK \tag{7.11.8}$$

Assume A transforms $|\psi\rangle$ and $|\phi\rangle$ into respectively

$$|\psi_A\rangle = A|\psi\rangle, \quad |\phi_A\rangle = A|\phi\rangle \tag{7.11.9}$$

then

$$\begin{aligned} \langle \phi_A \mid \psi_A\rangle &= \langle \phi| A^+ A |\psi\rangle = \langle\phi| K^+U^+UK|\psi\rangle \\ &= \langle \phi^* \mid \psi^*\rangle = \langle \phi \mid \psi\rangle^* = \langle \psi \mid \phi\rangle \end{aligned} \tag{7.11.10}$$

In the above operation K^+ only takes complex conjugation to $\langle\phi|$ on the left-hand side. This is consistent with

$$\langle\phi_A| = |\phi_A\rangle^+ = [UK|\phi\rangle]^+ = [U|\phi^*\rangle]^+ = \langle\phi^*| U^+$$

It is easy to see that

$$K^2 = 1, \quad K^{-1} = K, \quad A^{-1} = K^{-1}U^{-1} = K^{-1}U^+ \tag{7.11.11}$$

For any operator O, we can define

$$O_A = AOA^{-1} = UKOKU^+ = UO^*U^+ \tag{7.11.12}$$

So we can derive that

$$\langle\phi_A| O_A |\psi_A\rangle = \langle\phi| O |\psi\rangle^* \tag{7.11.13}$$

This equation will be used quite often.

Now we can discuss the constrains to M_{ij}, Γ_{ij} put by CPT, CP, T transformations. We first define the transformations for the state vectors.

Choose phase to make $(CP)^2 = 1$, i.e. state vectors do not change after two times of CP transformations. So we can define that

$$\begin{aligned} CP\left|\mathrm{P}^0\right\rangle &= \mathrm{e}^{\mathrm{i}\xi}\left|\bar{\mathrm{P}}^0\right\rangle \\ CP\left|\bar{\mathrm{P}}^0\right\rangle &= \mathrm{e}^{-\mathrm{i}\xi}\left|\mathrm{P}^0\right\rangle \end{aligned} \tag{7.11.14}$$

In the same way, we can define that

$$\begin{aligned} CPT\left|\mathrm{P}^0\right\rangle &= \mathrm{e}^{\mathrm{i}\nu}\left|\bar{\mathrm{P}}^0\right\rangle \\ CPT\left|\bar{\mathrm{P}}^0\right\rangle &= \mathrm{e}^{\mathrm{i}\nu}\left|\mathrm{P}^0\right\rangle \end{aligned} \tag{7.11.15}$$

Making CP operation on two sides of Eq. (7.11.15) and using Eq. (7.11.14), we have

$$\begin{aligned} T\left|\mathrm{P}^0\right\rangle &= \mathrm{e}^{\mathrm{i}(\nu-\xi)}\left|\mathrm{P}^0\right\rangle \\ T\left|\bar{\mathrm{P}}^0\right\rangle &= \mathrm{e}^{\mathrm{i}(\nu+\xi)}\left|\bar{\mathrm{P}}^0\right\rangle \end{aligned} \tag{7.11.16}$$

The phases of ν, ξ above are arbitrary. This does not affect any physical measurements. If weak interactions are invariant under the CPT, CP, T transformations, then we have

$$\begin{aligned} &(CPT)H_{\mathrm{W}}(CPT)^{-1} = H_{\mathrm{W}} \\ &(CP)H_{\mathrm{W}}(CP)^{+} = H_{\mathrm{W}} \\ &TH_{\mathrm{W}}T^{-1} = H_{\mathrm{W}} \end{aligned} \tag{7.11.17}$$

Using Eqs. (7.11.7) and (7.11.13)–(7.11.17) it is easy to prove the following constrains to M_{ij}, Γ_{ij} put by the invariance of weak interactions under CPT, CP, T transformations. We list the results in Table 7.1.

From Table 7.1, we can see that if interactions are invariant under T and CP transformations, then under T and CP transformations

$$M_{12}^*\Gamma_{12} = (\mathrm{e}^{-2\mathrm{i}\xi}M_{21})^*\mathrm{e}^{-2\mathrm{i}\xi}\Gamma_{21} = M_{21}^*\Gamma_{21} = M_{12}\Gamma_{12}^* \tag{7.11.18}$$

Table 7.1. The constrains of discrete symmetries to M_{ij}, Γ_{ij}.

Symmetry	Diagonal element	Non-diagonal element
CPT	$\Gamma_{11} = \Gamma_{22}$, $M_{11} = M_{22}$	no constrain
CP	$\Gamma_{11} = \Gamma_{22}$, $M_{11} = M_{22}$	$M_{12} = \mathrm{e}^{-2\mathrm{i}\xi}M_{21}$, $\Gamma_{12} = \mathrm{e}^{-2\mathrm{i}\xi}\Gamma_{21}$
T	no constrain	$M_{12} = \mathrm{e}^{-2\mathrm{i}\xi}M_{21}$, $\Gamma_{12} = \mathrm{e}^{-2\mathrm{i}\xi}\Gamma_{21}$

In the last step we used the Hermitian property of M and Γ, i.e. $M_{12} = M_{21}^*$, $\Gamma_{12} = \Gamma_{21}^*$, so we have

$$\mathrm{Im}(M_{12}^*\Gamma_{12}) = 0 \tag{7.11.19}$$

Therefore $\mathrm{Im}(M_{12}^*\Gamma_{12}) \neq 0$ means CP or T violation.

We denote the mass eigenstate of the effective Hamiltonian in Eq. (7.11.6) by

$$\begin{aligned} |\mathrm{P_H}\rangle &= p_\mathrm{H} \left|\mathrm{P}^0\right\rangle - q_\mathrm{H} \left|\bar{\mathrm{P}}^0\right\rangle \\ |\mathrm{P_L}\rangle &= p_\mathrm{L} \left|\mathrm{P}^0\right\rangle + q_\mathrm{L} \left|\bar{\mathrm{P}}^0\right\rangle \end{aligned} \tag{7.11.20}$$

where the subscripts H and L represent the mass heavy (H) and mass light (L) cases, respectively.

In general, either p_H and p_L or q_H and q_L are not equal. If we assume that the system is invariant under CPT transformation, then the problem is much more simplified. As for the general case the readers can refer [20], we do not discuss it here. Assuming CPT invariance, we can prove that [20] $p_\mathrm{H} = p_\mathrm{L} = p$, $q_\mathrm{H} = q_\mathrm{L} = q$. Equation (7.11.20) is reduced to

$$\begin{aligned} |\mathrm{P_H}\rangle &= p \left|\mathrm{P}^0\right\rangle - q \left|\bar{\mathrm{P}}^0\right\rangle \\ |\mathrm{P_L}\rangle &= p \left|\mathrm{P}^0\right\rangle + q \left|\bar{\mathrm{P}}^0\right\rangle \end{aligned} \tag{7.11.21}$$

The two corresponding eigenvalues of the two eigenstates are denoted by

$$\begin{aligned} \mu_\mathrm{H} &= m_\mathrm{H} - \frac{\mathrm{i}}{2}\Gamma_\mathrm{H} \\ \mu_\mathrm{L} &= m_\mathrm{L} - \frac{\mathrm{i}}{2}\Gamma_\mathrm{L} \end{aligned} \tag{7.11.22}$$

Define

$$\begin{aligned} \Delta m &= m_\mathrm{H} - m_\mathrm{L} \\ \Delta\Gamma &= \Gamma_\mathrm{H} - \Gamma_\mathrm{L} \end{aligned} \tag{7.11.23}$$

then the variation with time for $|\mathrm{P_H}\rangle$,$|\mathrm{P_L}\rangle$ is

$$\begin{aligned} |\mathrm{P_H}(t)\rangle &= \mathrm{e}^{-\mathrm{i}\mu_\mathrm{H} t} = \mathrm{e}^{-\mathrm{i}m_\mathrm{H} t}\mathrm{e}^{-\Gamma_\mathrm{H} t/2} \left|\mathrm{P}_H\right\rangle \\ |\mathrm{P_L}(t)\rangle &= \mathrm{e}^{-\mathrm{i}\mu_\mathrm{L} t} = \mathrm{e}^{-\mathrm{i}m_\mathrm{L} t}\mathrm{e}^{-\Gamma_\mathrm{L} t/2} \left|\mathrm{P_L}\right\rangle \end{aligned} \tag{7.11.24}$$

where t is the proper time in the rest frame of P^0, $\bar{\mathrm{P}}^0$. From Eqs. (7.11.21) and (7.11.24) we can derive

$$\begin{aligned} \left|\mathrm{P}^0(t)\right\rangle &= g_+(t)\left|\mathrm{P}^0\right\rangle + \frac{q}{p}g_-(t)\left|\bar{\mathrm{P}}^0\right\rangle \\ \left|\bar{\mathrm{P}}^0(t)\right\rangle &= \frac{p}{q}g_-(t)\left|\mathrm{P}^0\right\rangle + g_+(t)\left|\bar{\mathrm{P}}^0\right\rangle \end{aligned} \tag{7.11.25}$$

where

$$\begin{aligned} g_{\pm}(t) &= \frac{1}{2}\left\{ \mathrm{e}^{-\mathrm{i}\mu_{\mathrm{L}} t} \pm \mathrm{e}^{-\mathrm{i}\mu_{\mathrm{H}} t} \right\} \\ &= \frac{1}{2}\mathrm{e}^{-imt}\mathrm{e}^{-\Gamma t/2}\left\{ \mathrm{e}^{\mathrm{i}t\Delta m/2}\mathrm{e}^{\Delta\Gamma t/4} \pm \mathrm{e}^{-\mathrm{i}t\Delta m/2}\mathrm{e}^{-\Delta\Gamma t/4} \right\} \\ m &= (m_{\mathrm{H}} + m_{\mathrm{L}})/2 \end{aligned} \tag{7.11.26}$$

Now we compute q/p. For this reason we first solve the eigen-equation of Eq. (7.11.6), i.e.

$$\begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix} = \mu \begin{pmatrix} p \\ q \end{pmatrix} \tag{7.11.27}$$

It is easy to find that the eigenvalues are

$$\mu = \frac{H_{11} + H_{22} \pm \sqrt{4H_{12}H_{21} + (H_{11} - H_{22})^2}}{2} \tag{7.11.28}$$

The above two eigenvalues are denoted as

$$\begin{cases} \mu_{\mathrm{H}} = \frac{1}{2}(H_{11} + H_{22} + \sqrt{4H_{12}H_{21} + (H_{11} - H_{22})^2}) \\ \mu_{\mathrm{L}} = \frac{1}{2}(H_{11} + H_{12} - \sqrt{4H_{12}H_{21} + (H_{11} - H_{22})^2}) \end{cases} \tag{7.11.29}$$

where μ_{H} corresponds to the eigenstate $|\mathrm{P_H}\rangle$, and μ_{L} corresponds to $|\mathrm{P_L}\rangle$. Substituting them into the eigen-equation (7.11.27), we have

$$\begin{aligned} (H_{11} - \mu_{\mathrm{H}})\, p - H_{12} q = 0 \\ (H_{11} - \mu_{\mathrm{L}})\, p + H_{12} q = 0 \end{aligned} \tag{7.11.30}$$

So we get

$$\frac{q}{p} = -\frac{\mu_{\mathrm{H}} - \mu_{\mathrm{L}}}{2H_{12}} = -\frac{\sqrt{H_{12}H_{21}}}{H_{12}} = \pm\sqrt{\frac{H_{21}}{H_{12}}} \tag{7.11.31}$$

where $H_{11} = H_{22}$ (see Table 7.1) have been used for the case of *CPT* invariance. If we use m_{ij}, Γ_{ij} we can write

$$\frac{q}{p} = \pm\sqrt{\frac{2M_{12}^* - \mathrm{i}\Gamma_{12}^*}{2M_{12} - \mathrm{i}\Gamma_{21}}} \tag{7.11.32}$$

In Eq. (7.11.32) the signs + or − are related to the CP phases taken. Assuming CPT and CP invariance, from Table 7.1 we can obtain

$$\begin{aligned} M_{12}^* &= \mathrm{e}^{2\mathrm{i}\xi} M_{12} \\ \Gamma_{12}^* &= \mathrm{e}^{2\mathrm{i}\xi} \Gamma_{12} \end{aligned} \tag{7.11.33}$$

Substituting Eq. (7.11.33) into Eq. (7.11.32) we get

$$\frac{q}{p} = \pm \mathrm{e}^{\mathrm{i}\xi} \tag{7.11.34}$$

From Eq. (7.11.14) we see that $CP\left|\mathrm{P_H}\right\rangle = p\mathrm{e}^{\mathrm{i}\xi}\left|\bar{\mathrm{P}}^0\right\rangle - q\mathrm{e}^{-\mathrm{i}\xi}\left|\mathrm{P}^0\right\rangle$. Substituting $p = \pm q\mathrm{e}^{-\mathrm{i}\xi}$, $q = \pm \mathrm{e}^{\mathrm{i}\xi} p$ in Eq. (7.11.34) into the above equation, we get

$$\begin{aligned} CP\left|\mathrm{P_H}\right\rangle &= \pm q\mathrm{e}^{-\mathrm{i}\xi}\mathrm{e}^{\mathrm{i}\xi}\left|\bar{\mathrm{P}}^0\right\rangle \mp p\mathrm{e}^{\mathrm{i}\xi}\mathrm{e}^{-\mathrm{i}\xi}\left|\mathrm{P}^0\right\rangle \\ &= \mp\left\{p\left|\mathrm{P}^0\right\rangle - q\left|\bar{\mathrm{P}}^0\right\rangle\right\} = \mp\left|\mathrm{P_H}\right\rangle \end{aligned}$$

Thus

$$\begin{aligned} CP\left|\mathrm{P_H}\right\rangle &= \mp\left|\mathrm{P_H}\right\rangle \\ CP\left|\mathrm{P_L}\right\rangle &= \pm\left|\mathrm{P_L}\right\rangle \end{aligned} \tag{7.11.35}$$

Choose phase to make $CP\left|\mathrm{P_H}\right\rangle = -\left|\mathrm{P_H}\right\rangle$ (when CP conserves), then

$$\frac{q}{p} = \mathrm{e}^{\mathrm{i}\xi} \tag{7.11.36}$$

That is, the CP parity of $\left|\mathrm{P_H}\right\rangle$ is negative when CP conserves. Actually we take $\xi = 0$ later, i.e. we take

$$\begin{aligned} CP\left|\mathrm{P}^0\right\rangle &= \left|\bar{\mathrm{P}}^0\right\rangle \\ CP\left|\bar{\mathrm{P}}^0\right\rangle &= \left|\mathrm{P}^0\right\rangle \end{aligned} \tag{7.11.37}$$

Obviously Eqs. (7.11.36) and (7.11.37) guarantee the correctness of Eq. (7.11.35) when CP conserves. So we always take from now on

$$\frac{q}{p} = \sqrt{\frac{2M_{12}^* - \mathrm{i}\Gamma_{12}^*}{2M_{12} - \mathrm{i}\Gamma_{21}}} \tag{7.11.38}$$

In what follows, we give the classification of CP violation.

CP violations can be divided into three classes.

(1) Direct CP violation, or CP violation in decay amplitudes. If

$$|A\,(\mathrm{P} \to \mathrm{f})| \neq \left|\bar{A}\left(\bar{\mathrm{P}} \to \bar{\mathrm{f}}\right)\right| \tag{7.11.39}$$

then it is called direct CP violation, where P may be P^0 or $\mathrm{P}^{\pm}$, $\bar{\mathrm{f}}$ and $\bar{\mathrm{P}}$ are the CP conjugate states of the f and P, respectively. The partial decay rate asymmetry of the charged mesons $\mathrm{B}^{\pm}$ belongs to this category.

(2) Indirect CP violation, or CP violation in mixing. This is a kind of CP violation induced by neutral P^0–$\bar{\mathrm{P}}^0$ mixing, the condition is

$$\left|\frac{q}{p}\right| \neq 1 \tag{7.11.40}$$

(3) Mixed CP violation, or CP violation in the interference of mixing and decay. The condition is

$$\lambda = \frac{q}{p}\frac{\left|A\left(\bar{\mathrm{P}} \to \mathrm{f}\right)\right|}{|A\,(\mathrm{P} \to \mathrm{f})|} \neq 1 \tag{7.11.41}$$

Note that the case for $|\lambda| = 1$ but $\mathrm{Im}\,\lambda \neq 0$ is also belong to this category, and is very important special case. Later on we shall often encounter the above three classes of CP violation.

7.11.2. $\mathbf{K^0}$–$\mathbf{\bar{K}^0}$ *mixing and CP violation*

7.11.2.1. K^0–$\bar{\mathrm{K}}^0$ *mixing*

In the 1950s of 20 century, a series of strange particles, such as Λ^0, Ξ^-, were discovered. In order to explain creation and decay of the strange particles, Gell-Mann [31] and Nakano Nishijima [32] proposed a concept that the strange quantum number "S" was an additive quantum number. For K^-, Λ, $\bar{\mathrm{K}}^0$ their strange quantum number is -1, for K^+, $\bar{\Lambda}$ their strange quantum number is $+1$, while for Ξ^- its strange quantum number is -2, for all the other non-strange particles their strange quantum number is zero, and they assumed that the strange quantum number conserved in all strong interaction processes.

In experiments we can use $\pi^{\pm}$ beam to bombard nuclear target to produce K^0, $\bar{\mathrm{K}}^0$, e.g.

$$\begin{array}{lcccccc} & \pi^- & p & \to & \Lambda & \mathrm{K}^0 & \\ \text{strangeness S:} & 0 & 0 & & -1 & 1 & \\ & \pi^+ & \mathrm{p} & \to & \mathrm{K}^+ & \bar{\mathrm{K}}^0 & \mathrm{p} \\ \text{strangeness S:} & 0 & 0 & & 1 & -1 & 0 \end{array} \tag{7.11.42}$$

In the above processes, the strangeness conserves. Using quark language, $\mathrm{K}^0 = (\bar{\mathrm{s}}\mathrm{d})$, $K^+ = (\bar{\mathrm{s}}\mathrm{u})$, $\Lambda = (\mathrm{uds})$. So we only need to give -1 of strangeness to the s quark, we can explain the strangeness conservation of the processes in Eq. (7.11.42).

K^0, $\bar{\mathrm{K}}^0$ produced in the above strong interaction processes all can decay into 2π and 3π states, so K^0, $\bar{\mathrm{K}}^0$ have common decay final states therefore can transform into each other through 2π and 3π intermediate states, this is illustrated in Fig. 7.5.

The phenomenon of the K^0–$\bar{\mathrm{K}}^0$ mutual transformation is called K^0–$\bar{\mathrm{K}}^0$ mixing or K^0–$\bar{\mathrm{K}}^0$ oscillation. This subtle characteristic exists in other mesons, e.g. D^0–$\bar{\mathrm{D}}^0$, $\mathrm{B}^0_{\mathrm{d}}$–$\bar{\mathrm{B}}^0_{\mathrm{d}}$, $\mathrm{B}^0_{\mathrm{s}}$–$\bar{\mathrm{B}}^0_{s}$.

For K^0–$\bar{\mathrm{K}}^0$ system choose the CP phase in Eq. (7.11.14) $\xi = 0$ (the choice of ξ does not affect physically observable quantities), then

$$CP\left|\mathrm{K}^0\right\rangle = \left|\bar{\mathrm{K}}^0\right\rangle, \quad CP\left|\bar{\mathrm{K}}^0\right\rangle = \left|\mathrm{K}^0\right\rangle \tag{7.11.43}$$

From Eq. (7.11.21), we can define

$$\begin{aligned} |\mathrm{K_L}\rangle &= p\left|\mathrm{K}^0\right\rangle - q\left|\bar{\mathrm{K}}^0\right\rangle \\ |\mathrm{K_S}\rangle &= p\left|\mathrm{K}^0\right\rangle + q\left|\bar{\mathrm{K}}^0\right\rangle \end{aligned} \tag{7.11.44}$$

Here it is different from Eq. (7.11.21), In Eq. (7.11.44), we do not use subscripts of heavy (H) and light (L) to define mass eigenstates, but we use subscripts of long lifetime (L) and short lifetime (S) to define mass eigenstates for neutral K masons. This new definition is chosen according

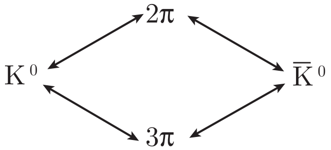


Fig. 7.5. Mutual transformation of K^0–$\bar{\mathrm{K}}^0$.

to the measured results in experiments. It is found in experiments that the lifetime of $\mathrm{K_L}$ is almost 570 times longer than that of $\mathrm{K_S}$. So for neutral K mesons it is more convenient to mark mass eigenstates by long and short lifetimes. As for differentiating them by heavy and light, indeed the mass of $\mathrm{K_L}$ is a little heavier than that of $\mathrm{K_S}$. It is measured in experiments [33]

$$\Delta m = m_{\mathrm{K_L}} - m_{\mathrm{K_S}} = (3.484 \pm 0.006) \times 10^{-12}\,\mathrm{MeV}$$

If we neglect the CP violation effect (this effect is very small, $\leqslant 10^{-3}$), then we can define

$$\begin{aligned} \left|\mathrm{K}_1^0\right\rangle &= \frac{1}{\sqrt{2}}\left(\left|\mathrm{K}^0\right\rangle + \left|\bar{\mathrm{K}}^0\right\rangle\right), CP = + \\ \left|\mathrm{K}_2^0\right\rangle &= \frac{1}{\sqrt{2}}\left(\left|\mathrm{K}^0\right\rangle - \left|\bar{\mathrm{K}}^0\right\rangle\right), CP = - \end{aligned} \tag{7.11.45}$$

From Eq. (7.11.43) we see that the CP parity of $\left|\mathrm{K}_1^0\right\rangle$ is positive, and the CP parity of $\left|\mathrm{K}_2^0\right\rangle$ is negative. Now we shall prove, for 2π states $\pi^+\pi^-$ and $\pi^0\pi^0$, no matter how much their relative momentum is, their CP parities are always positive, i.e.

$$CP\left|\pi^+\pi^-\right\rangle = \left|\pi^+\pi^-\right\rangle, \quad CP\left|\pi^0\pi^0\right\rangle = \left|\pi^0\pi^0\right\rangle \tag{7.11.46}$$

In fact, CP acting on $|\pi^+\pi^-\rangle$ is equivalent to exchange the position of $\pi^+\pi^-$ twice, so the appearing CP parity is $(-1)^{2l}(-1)^2 = 1$. For $\left|\pi^0\pi^0\right\rangle$, the total symmetry of the wave function for identical particle system demands that the relative angular momentum of the system l must be a even number, also because $CP\left|\pi^0\right\rangle = -\left|\pi^0\right\rangle$, thus $CP\left|\pi^0\pi^0\right\rangle = (-1)^l\left|\pi^0\pi^0\right\rangle = \left|\pi^0\pi^0\right\rangle$, its CP parity is also positive. Therefore we have proved that for 2π states their CP parities are always positive. So we see that for $\left|\mathrm{K}_1^0\right\rangle$ its CP parity is positive. The CP conservation requires that it can only decay into 2π states. While the CP parity of $\left|\mathrm{K}_2^0\right\rangle$ is negative, so it can only decays into 3π state with negative CP parity. The problem of CP parity of 3π states is a little bit complicated, we shall discuss it later.

It is found in experiments that $\mathrm{K_S}$ only decays to 2π, $\mathrm{K_L}$ only decays to 3π. So, when neglecting small CP violation effects,

$$|\mathrm{K_S}\rangle \approx \left|\mathrm{K}_1^0\right\rangle, \quad |\mathrm{K_L}\rangle \approx \left|\mathrm{K}_2^0\right\rangle \tag{7.11.47}$$

From Eq. (7.11.24), the time variation for $\mathrm{K_S}$, $\mathrm{K_L}$ is determined by Schrödinger equation, i.e.

$$\begin{aligned} \left|\mathrm{K_S}(t)\right\rangle &= \mathrm{e}^{-\mathrm{i}\left(m_{\mathrm{S}} - \frac{\mathrm{i}}{2}\Gamma_{\mathrm{S}}\right)t}\left|\mathrm{K_S}\right\rangle \\ \left|\mathrm{K_L}(t)\right\rangle &= \mathrm{e}^{-\mathrm{i}\left(m_{\mathrm{L}} - \frac{\mathrm{i}}{2}\Gamma_{\mathrm{L}}\right)t}\left|\mathrm{K_L}\right\rangle \end{aligned} \tag{7.11.48}$$

From that we can derive(see Eq. (7.11.25))

$$\begin{aligned}\left|\mathrm{K}^0(t)\right\rangle &= g_+(t)\left|\mathrm{K}^0\right\rangle + \frac{q}{p}g_-(t)\left|\bar{\mathrm{K}}^0\right\rangle \\ \left|\bar{\mathrm{K}}^0(t)\right\rangle &= \frac{p}{q}g_-(t)\left|\mathrm{K}^0\right\rangle + g_+(t)\left|\bar{\mathrm{K}}^0\right\rangle\end{aligned} \tag{7.11.49}$$

$$g_\pm(t) = \frac{1}{2}\mathrm{e}^{-\mathrm{i}mt}\mathrm{e}^{-\Gamma t/2}\{\mathrm{e}^{\mathrm{i}\Delta mt/2}\mathrm{e}^{\Delta\Gamma t/4} \pm \mathrm{e}^{-\mathrm{i}\Delta mt/2}\mathrm{e}^{-\Delta\Gamma t/4}\} \tag{7.11.50}$$

where

$$\begin{aligned}\Delta m &= m_{\mathrm{K_L}} - m_{\mathrm{K_S}} > 0 \\ \Delta\Gamma &= \Gamma_{\mathrm{K_L}} - \Gamma_{\mathrm{K_S}} < 0\end{aligned} \tag{7.11.51}$$

If at the initial time $t = 0$ the beam intensity of pure K^0 is I_0, then at time t the intensities of K^0 and $\bar{\mathrm{K}}^0$ found in the beam are respectively

$$\begin{aligned}I\left(\mathrm{K}^0 \to \mathrm{K}^0, t\right) &= I_0\left|\left\langle \mathrm{K}^0 \middle| \mathrm{K}^0(t)\right\rangle\right|^2 = I_0\left|g_+(t)\right|^2 \\ &= \frac{1}{4}I_0\{\mathrm{e}^{-\Gamma_{\mathrm{S}}t} + \mathrm{e}^{-\Gamma_{\mathrm{L}}t} + 2\mathrm{e}^{-\frac{1}{2}(\Gamma_{\mathrm{S}}+\Gamma_{\mathrm{L}})t}\cos\Delta m_{\mathrm{K}}t\}\end{aligned} \tag{7.11.52}$$

$$\begin{aligned}I\left(\mathrm{K}^0 \to \bar{\mathrm{K}}^0, t\right) &= I_0\left|\left\langle \bar{\mathrm{K}}^0 \middle| \mathrm{K}^0(t)\right\rangle\right|^2 = I_0\left|g_-(t)\right|^2 \\ &= \frac{1}{4}I_0\{\mathrm{e}^{-\Gamma_{\mathrm{S}}t} + \mathrm{e}^{-\Gamma_{\mathrm{L}}t} - 2\mathrm{e}^{-\frac{1}{2}(\Gamma_{\mathrm{S}}+\Gamma_{\mathrm{L}})t}\cos\Delta m_{\mathrm{K}}t\}\end{aligned} \tag{7.11.53}$$

Obviously, the intensities of K^0, $\bar{\mathrm{K}}^0$ all oscillate with frequency Δm_{K}, the oscillating curves are shown in Fig. 7.6. Where the intensities of K^0 and

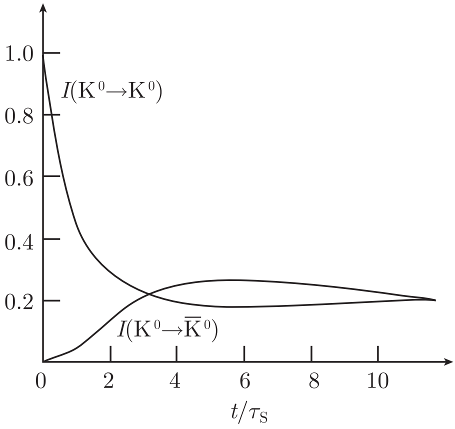


Fig. 7.6.

$\bar{\mathrm{K}}^0$ are normalized by I_0 already, and $\Delta m_{\mathrm{K}}\tau_{\mathrm{S}} = 0.5$ is taken, τ_{S} is the K_{S} average lifetime. The time t in Eqs. (7.11.52) and (7.11.53) is the time at the K^0 rest frame,

$$t = d/\beta\gamma \tag{7.11.54}$$

where d is the flying distance of K^0 in the laboratory, $\beta = v/c$, $\gamma = 1/\sqrt{1-\beta^2}$, v is the flying velocity of K^0 in the laboratory frame. The above discussion is proceeded in neglecting CP violation effects. That is, K^0–$\bar{\mathrm{K}}^0$ oscillation is irrelevant to the CP violation of K^0–$\bar{\mathrm{K}}^0$ system.

7.11.2.2. *Reproduction of neutral* K *meson*

From Eqs. (7.11.45) and (7.11.47), if we neglect CP violation, then

$$\begin{aligned} \left|\mathrm{K}^0_{\mathrm{S}}\right\rangle &= \frac{1}{\sqrt{2}}\left(\left|\mathrm{K}^0\right\rangle + \left|\bar{\mathrm{K}}^0\right\rangle\right) \\ \left|\mathrm{K}^0_{\mathrm{L}}\right\rangle &= \frac{1}{\sqrt{2}}\left(\left|\mathrm{K}^0\right\rangle - \left|\bar{\mathrm{K}}^0\right\rangle\right) \end{aligned} \tag{7.11.55}$$

Obviously

$$\begin{aligned} \left|\mathrm{K}^0\right\rangle &= \frac{1}{\sqrt{2}}\left(\left|\mathrm{K}_{\mathrm{S}}\right\rangle + \left|\mathrm{K}_{\mathrm{L}}\right\rangle\right) \\ \left|\bar{\mathrm{K}}^0\right\rangle &= \frac{1}{\sqrt{2}}\left(\left|\mathrm{K}_{\mathrm{S}}\right\rangle - \left|\mathrm{K}_{\mathrm{L}}\right\rangle\right) \end{aligned} \tag{7.11.56}$$

In experiments, if we use $\pi^-\mathrm{p} \to \Lambda\mathrm{K}^0$ to produce pure K^0 beam, what will happen when time goes on? Because in K^0 beam a half is K_{S}, another half is K_{L}. After flying 2.68 cm on average, K_{S} will disappear through decaying into 2π. While K_{L} can fly 15.34 m on average before its decay. So from the K^0 production point about ten meters away only K_{L} is left. But from Eq. (7.11.55) we see that in K_{L} a half is K^0, another half is $\bar{\mathrm{K}}^0$, so a pure K^0 beam at initial time, after more than ten meters flying, a half $\bar{\mathrm{K}}^0$ mixes in owing to the oscillation induced by weak interactions.

If we put a metal target on the flying path of K_{L}, because K^0 only has elastic and semi-elastic scatterings of strong interactions in nuclear media, but $\bar{\mathrm{K}}^0$ in addition to the above scatterings, there are inelastic interaction processes, for instance, the productions of Σ and Λ, so $\bar{\mathrm{K}}^0$ is absorbed easier than K^0. If the metal target is thick enough, then after K_{L} penetrating it, in $\left|\mathrm{K}^0_{\mathrm{L}}\right\rangle = \frac{1}{\sqrt{2}}\left(\left|\mathrm{K}^0\right\rangle - \left|\bar{\mathrm{K}}^0\right\rangle\right)$ only K^0 component is left, while $\left|\mathrm{K}^0\right\rangle = \frac{1}{\sqrt{2}}\left(\left|\mathrm{K}_{\mathrm{S}}\right\rangle + \left|\mathrm{K}_{\mathrm{L}}\right\rangle\right)$, so after K_{L} penetrates the metal target, K_{S} is reproduced. This subtle phenomenon happens because the lifetimes of K_{S}

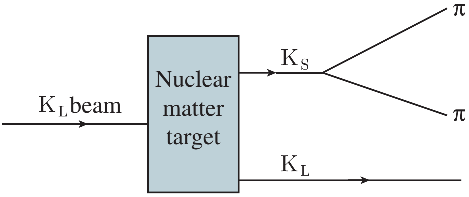


Fig. 7.7. Sketch of K_S reproduction.

and K_L have several hundred times difference and K^0 and $\bar{K}^0$ have different interactions with nuclear media. The reproduction of the K_S is illustrated in Fig. 7.7.

7.11.2.3. *Discovery of CP violation*

In Sec. 11.2.1, we have explained that when neglecting CP violation K_L cannot decay into 2π final state. This is because the CP parity of K_L is negative, while the CP parity of 2π state is positive.

In 1964, Cronin, Fitch, Christenson and Turly [10] firstly discovered in Brookhaven Laboratory the decay of $K_L \to 2\pi$ state. The double arm spectrometer used in this experiment is sketched in Fig. 7.8. In this experiment, the beam of K_S and K_L firstly flies 57 feet and then arrives the front of the spectrometer. In this way at the front of the spectrometer only K_L beam is left. K_L goes into the bag filled up with helium and produces decay of $K_L \to 2\pi$. The measured result is

$$R = \frac{\mathrm{Br}\,(K_L \to \pi^+\pi^-)}{\mathrm{Br}\,(K_L \to \text{all charged final states})} = (2.0 \pm 0.4) \times 10^{-3} \quad (7.11.57)$$

This result shows that CP symmetry is broken in the decay of $K_L \to 2\pi$. This is shown again that many symmetries taken as accurate before are no longer accurate, and are actually approximate. After P,C are proved not conserved, CP is also not conserved. So, K_L is not equivalent accurately to K_2^0, in K_L there must mix a small amount of K_1^0 component, i.e.

$$\begin{aligned} |K_L\rangle &= \frac{1}{\sqrt{1+|\bar{\varepsilon}|^2}}\left(|K_2^0\rangle + \bar{\varepsilon}\,|K_1^0\rangle\right) \\ |K_S\rangle &= \frac{1}{\sqrt{1+|\bar{\varepsilon}|^2}}\left(|K_1^0\rangle + \bar{\varepsilon}\,|K_2^0\rangle\right) \end{aligned} \quad (7.11.58)$$

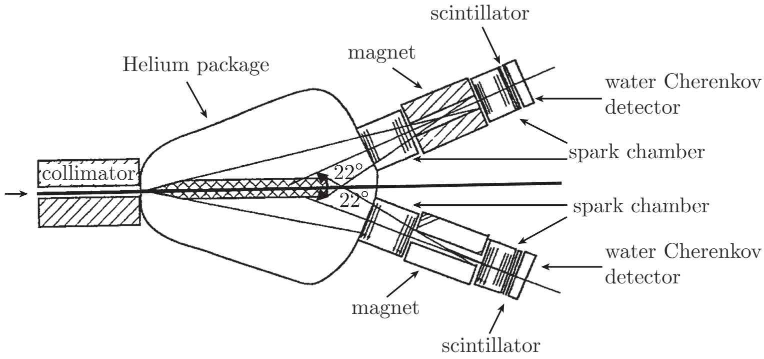

scintillator
magnet
Helium package
water Cherenkov detector
collimator
22°
22°
spark chamber
spark chamber
water Cherenkov detector
magnet
scintillator

Fig. 7.8. Sketch of the double arm spectrometer for discovering $K_L \to 2\pi$.

The experiment gives [10],

$$\bar{\varepsilon} = 2.3 \times 10^{-3} \tag{7.11.59}$$

so we can see that the strength of CP violation is very small, only at the order of one part in a thousand.

About 10 years later, in the semi-leptonic decays of K_L CP violation was also observed. It was found that the total width of the decay $K_L \to l^+\nu_l\pi^-$ was different with that of the decay into CP conjugate final state $K_L \to l^-\bar{\nu}_l\pi^+$. The CP violation parameter [34]

$$\begin{aligned} A_l &= \frac{\Gamma(K_L \to l^+\nu_l\pi^-) - \Gamma(K_L \to l^-\bar{\nu}_l\pi^+)}{\Gamma(K_L \to l^+\nu_l\pi^-) + \Gamma(K_L \to l^-\bar{\nu}_l\pi^+)} \\ &= (0.332 \pm 0.006)\,\% \end{aligned} \tag{7.11.60}$$

The final value is taken from the PDG book of the year 2012 [33].

7.11.2.4. *Isospin description of* $K^0 \to 2\pi$ *decays*

The final state of the decay $K^0 \to 2\pi$ is in s-wave, the symmetry of the identical bosons demands that the total isospin of 2π can be only 0, 2, so the decays of neutral K mesons to 2π can be described by the isospin amplitudes $A_I\,(I = 0, 2)$. Define

$$\begin{aligned} \langle (2\pi)_I | H_W | K^0 \rangle &= A_I e^{i\delta_I} \\ \langle (2\pi)_I | H_W | \bar{K}^0 \rangle &= \bar{A}_I e^{i\delta_I} \end{aligned} \tag{7.11.61}$$

where δ_I is the strong phase of the s wave for $(2\pi)_I$ state at energy $E = m_{\mathrm{K}^0}$, while A_I,$\bar{A}_I$ still include the phases of weak interactions. Using the method in Sec. 7.11.1 we can prove that [20]

$$CPT \text{ invariance leads to } A_I = \bar{A}_I^* \tag{7.11.62}$$

$$CP \text{ invariance leads to } A_I = \bar{A}_I \tag{7.11.63}$$

$$T \text{ invariance leads to } A_I = A_I^* \tag{7.11.64}$$

Because the π final states of K^0–$\bar{\mathrm{K}}^0$ decays can be only $\pi^+\pi^-$, $\pi^0\pi^0$, the total isospins of these final states can be only $(I_3 = 0, I = 0)$ and $(I_3 = 0, I = 2)$, from isospin-coupling analysis the states $(2\pi)_{II_3}$ can be expressed as

$$\begin{aligned} |(2\pi)_{0,0}\rangle &= \frac{1}{\sqrt{3}}\left|\pi^+\pi^-\right\rangle - \frac{1}{\sqrt{3}}\left|\pi^0\pi^0\right\rangle + \frac{1}{\sqrt{3}}\left|\pi^-\pi^+\right\rangle \\ |(2\pi)_{2,0}\rangle &= \frac{1}{\sqrt{6}}\left|\pi^+\pi^-\right\rangle + \sqrt{\frac{2}{3}}\left|\pi^0\pi^0\right\rangle + \frac{1}{\sqrt{6}}\left|\pi^-\pi^+\right\rangle \end{aligned} \tag{7.11.65}$$

In experiments we do not distinguish the states of $\pi^+\pi^-$ and $\pi^-\pi^+$, so we symmetrize them. Define

$$\begin{aligned} \left|\pi^+\pi^-\right\rangle_{\mathrm{S}} &= \frac{|\pi^+\pi^-\rangle + |\pi^-\pi^+\rangle}{\sqrt{2}} \\ &= \frac{|\pi^+(\vec{p})\,\pi^-(-\vec{p})\rangle + |\pi^-(\vec{p})\,\pi^+(-\vec{p})\rangle}{\sqrt{2}} \end{aligned} \tag{7.11.66}$$

Actually, later we shall neglect the subscript S, $|\pi^+\pi^-\rangle$ is understood as the state of $|\pi^+\pi^-\rangle_{\mathrm{S}}$ in Eq. (7.11.66). Equation (7.11.65) is simplified as

$$\begin{aligned} |(2\pi)_{0,0}\rangle &= \sqrt{\frac{2}{3}}\left|\pi^+\pi^-\right\rangle - \sqrt{\frac{1}{3}}\left|\pi^0\pi^0\right\rangle \\ |(2\pi)_{2,0}\rangle &= \sqrt{\frac{1}{3}}\left|\pi^+\pi^-\right\rangle + \sqrt{\frac{2}{3}}\left|\pi^0\pi^0\right\rangle \end{aligned} \tag{7.11.67}$$

From Eq. (7.11.67) we can derive

$$\begin{aligned} |\pi^+\pi^-\rangle &= \sqrt{\frac{2}{3}}|(2\pi)_{0,0}\rangle + \sqrt{\frac{1}{3}}|(2\pi)_{2,0}\rangle \\ |\pi^0\pi^0\rangle &= \sqrt{\frac{2}{3}}|(2\pi)_{2,0}\rangle - \sqrt{\frac{1}{3}}|(2\pi)_{0,0}\rangle \end{aligned} \tag{7.11.68}$$

According to Eqs. (7.11.61) and (7.11.68) we have

$$\begin{aligned}
\langle\pi^+\pi^-|\,H_{\mathrm{W}}|\mathrm{K}^0\rangle &= \sqrt{\frac{2}{3}}A_0\mathrm{e}^{\mathrm{i}\delta_0}+\sqrt{\frac{1}{3}}A_2\mathrm{e}^{\mathrm{i}\delta_2}\\
\left\langle\pi^0\pi^0\right|H_{\mathrm{W}}|\mathrm{K}^0\rangle &= \sqrt{\frac{2}{3}}A_2\mathrm{e}^{\mathrm{i}\delta_2}-\sqrt{\frac{1}{3}}A_0\mathrm{e}^{\mathrm{i}\delta_0}\\
\langle\pi^+\pi^-|\,H_{\mathrm{W}}\left|\bar{\mathrm{K}}^0\right\rangle &= \sqrt{\frac{2}{3}}\bar{A}_0\mathrm{e}^{\mathrm{i}\delta_0}+\sqrt{\frac{1}{3}}\bar{A}_2\mathrm{e}^{\mathrm{i}\delta_2}\\
\left\langle\pi^0\pi^0\right|H_{\mathrm{W}}\left|\bar{\mathrm{K}}^0\right\rangle &= \sqrt{\frac{2}{3}}\bar{A}_2\mathrm{e}^{\mathrm{i}\delta_2}-\sqrt{\frac{1}{3}}\bar{A}_0\mathrm{e}^{\mathrm{i}\delta_0}
\end{aligned} \tag{7.11.69}$$

Using Eq. (7.11.44), we obtain

$$\begin{aligned}
\langle\pi^+\pi^-|\,H_{\mathrm{W}}\,|\mathrm{K_L}\rangle &= \sqrt{\frac{2}{3}}\mathrm{e}^{\mathrm{i}\delta_0}pA_0\left(\Delta_0+\frac{1}{\sqrt{2}}\omega\Delta_2\mathrm{e}^{\mathrm{i}(\delta_2-\delta_0)}\right)\\
\left\langle\pi^0\pi^0\right|H_{\mathrm{W}}\,|\mathrm{K_L}\rangle &= \sqrt{\frac{2}{3}}\mathrm{e}^{\mathrm{i}\delta_0}pA_0\left(\omega\Delta_2\mathrm{e}^{\mathrm{i}(\delta_2-\delta_0)}-\frac{1}{\sqrt{2}}\Delta_0\right)
\end{aligned} \tag{7.11.70}$$

$$\begin{aligned}
\langle\pi^+\pi^-|\,H_{\mathrm{W}}\,|\mathrm{K_S}\rangle &= \sqrt{\frac{2}{3}}\mathrm{e}^{\mathrm{i}\delta_0}pA_0\left((2-\Delta_0)+\frac{1}{\sqrt{2}}\omega(2-\Delta_2)\mathrm{e}^{\mathrm{i}(\delta_2-\delta_0)}\right)\\
\left\langle\pi^0\pi^0\right|H_{\mathrm{W}}\,|\mathrm{K_S}\rangle &= \sqrt{\frac{2}{3}}\mathrm{e}^{\mathrm{i}\delta_0}pA_0\left(\omega(2-\Delta_2)\mathrm{e}^{\mathrm{i}(\delta_2-\delta_0)}-\frac{1}{\sqrt{2}}(2-\Delta_0)\right)
\end{aligned} \tag{7.11.71}$$

where

$$\Delta_I = 1-\frac{q}{p}\frac{\bar{A}_I}{A_I},\quad \omega=\frac{A_2}{A_0} \tag{7.11.72}$$

Define two experimental observables η_{+-},η_{00},

$$\eta_{+-}=\frac{\langle\pi^+\pi^-|\,H_{\mathrm{W}}\,|\mathrm{K_L}\rangle}{\langle\pi^+\pi^-|\,H_{\mathrm{W}}\,|\mathrm{K_S}\rangle},\quad \eta_{00}=\frac{\left\langle\pi^0\pi^0\right|H_{\mathrm{W}}\,|\mathrm{K_L}\rangle}{\langle\pi^0\pi^0|\,H_{\mathrm{W}}\,|\mathrm{K_S}\rangle} \tag{7.11.73}$$

Here we can still make approximation in the denominator of the above formulas. Because the probability of the decay of $\mathrm{K_S}\to\pi\pi$ is almost 100%, so in the calculation of $\langle\pi\pi|\,H_{\mathrm{W}}\,|\mathrm{K_S}\rangle$ CP violation effects can be neglected. So we can take $A_I=\bar{A}_I$, $q=p=1/\sqrt{2}$, thus $\Delta_0=\Delta_2=0$. Therefore we have

$$\begin{aligned}\langle\pi^+\pi^-|\,H_\mathrm{W}\,|\mathrm{K_S}\rangle &\approx \sqrt{\frac{2}{3}}\mathrm{e}^{\mathrm{i}\delta_0}pA_0\left(2+\omega\sqrt{2}\mathrm{e}^{\mathrm{i}(\delta_2-\delta_0)}\right)\\ \left\langle\pi^0\pi^0\right|H_\mathrm{W}\,|\mathrm{K_S}\rangle &\approx \sqrt{\frac{2}{3}}\mathrm{e}^{\mathrm{i}\delta_0}pA_0\left(2\omega\mathrm{e}^{\mathrm{i}(\delta_2-\delta_0)}-\sqrt{2}\right)\end{aligned}\tag{7.11.74}$$

Because the constrain of $|\Delta I|=1/2$ rule, $|\omega|=|A_2/A_0|\sim 1/20$, so in the computation of η_{+-},η_{00} we can neglect ω^2 term. So, from Eqs. (7.11.70), (7.11.73), and (7.11.74) we obtain

$$\begin{aligned}\eta_{+-} &\approx \frac{1}{2}\left(\Delta_0-\frac{1}{\sqrt{2}}\omega(\Delta_0-\Delta_2)\mathrm{e}^{\mathrm{i}(\delta_2-\delta_0)}\right)\\ \eta_{00} &\approx \frac{1}{2}\left(\Delta_0+\sqrt{2}\omega(\Delta_0-\Delta_2)\mathrm{e}^{\mathrm{i}(\delta_2-\delta_0)}\right)\end{aligned}\tag{7.11.75}$$

From η_{+-},η_{00} we can define experimentally measurable parameters ε and ε':

$$\begin{cases}\eta_{+-}=\varepsilon+\varepsilon'\\ \eta_{00}=\varepsilon-2\varepsilon'\end{cases}\Rightarrow\begin{cases}\varepsilon=\dfrac{2\eta_{+-}+\eta_{00}}{3}\\ \varepsilon'=\dfrac{\eta_{+-}-\eta_{00}}{3}\end{cases}\tag{7.11.76}$$

Substituting Eq. (7.11.75) into Eq. (7.11.76) we obtain

$$\begin{aligned}\varepsilon &= \frac{1}{2}\Delta_0\\ \varepsilon' &= -\frac{1}{2\sqrt{2}}\omega(\Delta_0-\Delta_2)\mathrm{e}^{\mathrm{i}(\delta_2-\delta_0)}\end{aligned}\tag{7.11.77}$$

η_{+-},η_{00} can all be measured quantities. $|\eta_{+-}|,|\eta_{00}|$ can be measured through measuring the branching ratios of $\mathrm{K_{L,S}}\to 2\pi$. These phases also can be measured, the detail method can be found in [20]. Certainly ε,ε' can be measured, here we list their values below [33] (assuming *CPT* invariance):

$$\begin{aligned}&|\eta_{00}|=(2.220\pm0.011)\times10^{-3},\quad \phi_{00}=(43.52\pm0.05)^\circ\\ &|\eta_{+-}|=(2.232\pm0.011)\times10^{-3},\quad \phi_{+-}=(43.51\pm0.05)^\circ\\ &|\varepsilon|=(2.228\pm0.011)\times10^{-3},\quad \phi_{\varepsilon}=(43.52\pm0.05)^\circ\\ &\mathrm{Re}\,(\varepsilon'/\varepsilon)=(1.66\pm0.23)\times10^{-3}\\ &\mathrm{Im}\,(\varepsilon'/\varepsilon)=-\frac{1}{3}\left(\phi_{+-}-\phi_{00}\right)=(-0.002\pm0.005)^\circ\end{aligned}$$

7.11.2.5. *CP violation in* $\mathrm{K}^0 \to 3\pi$ *decays*

The decays of the neutral K masons to 3π are quite complicated. The 3π final states with zero of total charge have two kinds: $3\pi^0$ and $\pi^+\pi^-\pi^0$.

For $\left|\pi^0\pi^0\pi^0\right\rangle$ state, because the total isospin of $\pi^0\pi^0$ can be only 0,2, so the total isospin of $3\pi^0$ state can only be 1 and 3. $\left|\pi^0\pi^0\pi^0\right\rangle$ state is symmetric under exchange of any two π^0, and $CP\left|\pi^0\right\rangle = -\left|\pi^0\right\rangle$, thus

$$CP\left|\pi^0\pi^0\pi^0\right\rangle = -\left|\pi^0\pi^0\pi^0\right\rangle \tag{7.11.78}$$

For the state of $\pi^+\pi^-\pi^0$, assuming the relative angular momentum of $\pi^+\pi^-$ is l, the relative momentum of π^0 and $\pi^+\pi^-$ is l'. Because for $\mathrm{K_L}$, $\mathrm{K_S}$, $J^{\mathrm{P}} = 0^-$, so the total momentum of the 3π final states must be zero, this demands $l = l'$. No matter how much is l, we always have (see Eq. (7.11.46))

$$CP\left|\pi^+\pi^-\right\rangle = \left|\pi^+\pi^-\right\rangle \tag{7.11.79}$$

After π^0 comes in, because $CP\left|\pi^0\right\rangle = -\left|\pi^0\right\rangle$, and the factor of $(-1)^{l'}$ appears, so

$$\begin{aligned} CP\left|\pi^+\pi^-\pi^0\right\rangle &= -(-1)^{l'}\left|\pi^+\pi^-\pi^0\right\rangle = -(-1)^{l}\left|\pi^+\pi^-\pi^0\right\rangle \\ &= (-1)^{l+1}\left|\pi^+\pi^-\pi^0\right\rangle \end{aligned} \tag{7.11.80}$$

Now we list the relation for $\pi^+\pi^-$ relative angular momentum l, $\pi^+\pi^- l\pi^0$ total isospin I and CP parity in Table 7.2.

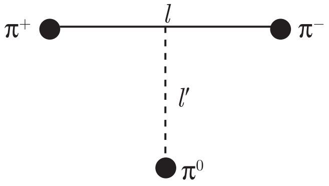


Fig. 7.9. The relative angular momentum of the 3π system.

Table 7.2. Relation for CP, l, I.

$\pi^+\pi^-$ isospin	0	1	2
$\pi^+\pi^-$ angular momentum l	even	odd	even
$\pi^+\pi^-\pi^0$ total isospin I	1	0, 2	1, 3
$\pi^+\pi^-\pi^0$ total CP parity	$-$	$+$	$-$

From Table 7.2, we see that the CP parity of 3π

$$\eta_{CP} = (-1)^I \tag{7.11.81}$$

For $3\pi^0$, CP parity is negative ($I = 1, 3$), while the CP parity of $\pi^+\pi^-\pi^0$ state can be either positive or negative. If the decay of $\mathrm{K_S} \to 3\pi^0$ is observed, then CP is violated. In experiments, we can measure

$$\eta_{000} = \frac{A(\mathrm{K_S} \to 3\pi^0)}{A(\mathrm{K_L} \to 3\pi^0)} \tag{7.11.82}$$

where A denotes the corresponding decay amplitude (cf. Eq. (7.1.73)). The PDG book of 2012 gives [33] (CL expresses confidence level),

$$|\eta_{000}| < 0.018, \quad \mathrm{CL} = 90\%$$

$$\mathrm{Im}\eta_{000} = (-0.1 \pm 1.6) \times 10^{-2}, \quad \mathrm{CL} = 90\%$$

For $\pi^+\pi^-\pi^0$ final states, two kinds of $CP = \pm$ states are mixed. From Table 7.2, we see that the total isospin of $CP = +$ state is 0 or 2, while the isospin of the corresponding $\pi^+\pi^-$ combination is 1, and the relative angular momentum is odd, so the exchange of $\pi^+\pi^-$ is antisymmetric. Thus, the total integrated contribution to $\pi^+\pi^-$ in Dalitz plot for $CP = +$ states of $\mathrm{K_S} \to \pi^+\pi^-\pi^0$ decays is zero. Therefore, we can see only CP negative states of $\mathrm{K_S} \to \pi^+\pi^-\pi^0$ using Dalitz plot method. But the final states of $CP = +$ in $\mathrm{K_L} \to \pi^+\pi^-\pi^0$ decays are suppressed owing to the angular momentum centrifugal barrier of $l = 1, 3, 5, \ldots$ for $\pi^+\pi^-$, so

$$\eta_{+-0} = \frac{A(\mathrm{K_S} \to \pi^+\pi^-\pi^0; CP = -1)}{A(\mathrm{K_L} \to \pi^+\pi^-\pi^0)} \tag{7.11.83}$$

also describes CP violation in the decays of $\mathrm{K_S} \to \pi^+\pi^-\pi^0$. The *PDG* book of 2012 gives [33]

$$\mathrm{Im}\eta_{+-0} = -0.002 \pm 0.009, \quad \mathrm{CL} = 90\%$$

7.11.2.6. *The relation of $\bar{\varepsilon}$ and ε*

In Eq. (7.11.58) $|\bar{\varepsilon}|$ represents the proportionality of K_1^0 mixed in $\mathrm{K_L}$, and we already measured the magnitude of $|\bar{\varepsilon}|$ in experiments. From Eq. (7.11.58) we can see that the phase of $\bar{\varepsilon}$ varies according to the phases of K_1^0, in other words to the phases of K^0, $\bar{\mathrm{K}}^0$, so it is not physical observable, this is different from η_{+-}, η_{00}. From Eq. (7.10.75) we see that η_{+-}, η_{00} depends on $\omega = \frac{A_2}{A_1}$ and $\Delta_I = 1 - \frac{q}{p}\frac{\bar{A}_I}{A_I}$. From Eq. (7.11.61), ω does not depend on the choice of the phases for K^0 so can be measured, $\dfrac{q}{p}\dfrac{\bar{A}_I}{A_I}$ also does not depend on the choice of the phases of K^0 and $\bar{\mathrm{K}}^0$. Actually, when

$$\left|\mathrm{K}^0\right\rangle \to \mathrm{e}^{\mathrm{i}\beta}\left|\mathrm{K}^0\right\rangle, \quad \left|\bar{\mathrm{K}}^0\right\rangle \to \mathrm{e}^{\mathrm{i}\bar{\beta}}\left|\bar{\mathrm{K}}^0\right\rangle$$

$$\frac{\bar{A}_I}{A_I} \to \mathrm{e}^{\mathrm{i}(\bar{\beta}-\beta)}\frac{\bar{A}_I}{A_I}, \quad \frac{q}{p} = \frac{\sqrt{M_{12}^* - \frac{\mathrm{i}}{2}\Gamma_{12}^*}}{\sqrt{M_{12} - \frac{\mathrm{i}}{2}\Gamma_{12}}} \to \mathrm{e}^{-\mathrm{i}(\bar{\beta}-\beta)}\frac{q}{p}$$

thus $\frac{q}{p}\frac{\bar{A}_I}{A_I}$ keeps unchanged. The strong phase difference $\delta_2 - \delta_0$ is certainly not changed, so η_{+-}, η_{00} and ε, ε' all keep invariant under the phase transformation of $\left|\mathrm{K}^0\right\rangle$, $\left|\bar{\mathrm{K}}^0\right\rangle$.

In Wu–Yang phase convention [35], A_0 is taken to be real, this can be done through the choice of $\left|\mathrm{K}^0\right\rangle$. On the condition of CPT invariance Eq. (7.11.62) gives

$$\bar{A}_0 = A_0^* = A_0 \tag{7.11.84}$$

So from Eq. (7.11.77) we have

$$\varepsilon = \frac{1}{2}\left(1 - \frac{q}{p}\frac{\bar{A}_0}{A_0}\right) = \frac{1}{2}\left(1 - \frac{q}{p}\right) = \frac{1}{2}\left(1 - \frac{1-\bar{\varepsilon}}{1+\bar{\varepsilon}}\right) \tag{7.11.85}$$

In the last step above the definition of $\bar{\varepsilon}$ in Eq. (7.11.58) is used.

Because $|\bar{\varepsilon}| \sim 10^{-3} \ll 1$, so in Eq. (7.11.85) we can take the following approximation

$$\varepsilon \approx \bar{\varepsilon} \tag{7.11.86}$$

This means that under Wu–Yang-phase convention, ε and $\bar{\varepsilon}$ are approximately equal; certainly we also have

$$|\varepsilon| \approx |\bar{\varepsilon}| \tag{7.11.87}$$

7.11.3. $\mathbf{B^0}$–$\mathbf{\bar{B}^0}$ *mixing and CP violation*

B mesons discovered in experiments include $\mathrm{B}_\mathrm{d}^0 = (\bar{\mathrm{b}}\mathrm{d})$, $\mathrm{B}_\mathrm{s}^0 = (\bar{\mathrm{b}}\mathrm{s})$, $\mathrm{B}_\mathrm{u}^+ = (\bar{\mathrm{b}}\mathrm{u})$, $\mathrm{B}_\mathrm{c}^+ = (\bar{\mathrm{b}}\mathrm{c})$ and their corresponding antiparticles. The masses of these B mesons are all around 5 GeV. In this subsection, we only discuss $\mathrm{B}_\mathrm{d}^0 - \bar{\mathrm{B}}_\mathrm{d}^0$ mixing and CP violation, for other B mesons the readers can consult the corresponding references.

7.11.3.1. *Experimental measurements of* B_d^0–$\bar{\mathrm{B}}_\mathrm{d}^0$ *mixing*

According to the standard electroweak unified model B_d^0 can only decay to positively charged leptons, $\bar{\mathrm{B}}_\mathrm{d}^0$ can only decay to negatively charged leptons. So the ways of the semi-leptonic decays for B_d^0, $\bar{\mathrm{B}}_\mathrm{d}^0$ are

$$\mathrm{B}_\mathrm{d}^0 \to l^+\nu_l\mathrm{X}, \quad \bar{\mathrm{B}}_\mathrm{d}^0 \to l^-\bar{\nu}_l\mathrm{X}$$

The bond state of $\mathrm{b}\bar{\mathrm{b}}$ $\Upsilon(9.5)$ was firstly discovered by Lederman's Group at Fermilaboratory 400 GeV proton synchrotron in 1977 [36]. The process they used is

$$\begin{tikzcd} \mathrm{p}+(\mathrm{Cu},\ \mathrm{Pt}) \arrow[r] & \Upsilon(9.5) \arrow[d] & +\mathrm{X} \\ & \mu^+\mu^- & \end{tikzcd}$$

Later, at German $\mathrm{e}^+\mathrm{e}^-$ collider (DESY) and Cornell University $\mathrm{e}^+\mathrm{e}^-$ collider (CERS) all discovered $\Upsilon(9.5)$. $\Upsilon(9.5)$ is $\mathrm{b}\bar{\mathrm{b}}$ $1S$ bound state, denoted by $\Upsilon(1S)$. Later on, the highly excited bound states of $\mathrm{b}\bar{\mathrm{b}}$ $\Upsilon(2S)$, $\Upsilon(3S)$, $\Upsilon(4S)$ etc. were successively discovered. All these states are $J^{PC} = 1^{--}$ vector mesons. They also found at experiments that $\Upsilon(4S)$ decayed almost 100% into $\mathrm{B}_\mathrm{d}^0\bar{\mathrm{B}}_\mathrm{d}^0+\mathrm{B}_\mathrm{u}^+\mathrm{B}_\mathrm{u}^-$, where $\mathrm{B}_\mathrm{d}^0\bar{\mathrm{B}}_\mathrm{d}^0$ and $\mathrm{B}_\mathrm{u}^+\mathrm{B}_\mathrm{u}^-$ are 50% each, respectively.

In 1987, at Germany Hamburg $\mathrm{e}^+\mathrm{e}^-$ collider(DESY), ARGUS Collaboration Group firstly observed the same sign leptonic events in final states $\mathrm{B}_\mathrm{d}^0\bar{\mathrm{B}}_\mathrm{d}^0 \to \mu^\pm\mu^\pm + \mathrm{X}$ [37]. This means that $\mathrm{B}_\mathrm{d}^0 \leftrightarrow \bar{\mathrm{B}}_\mathrm{d}^0$ oscillation (mixing) must happen, or else the same sign leptons could not appear in the final states. From these, they measured the mixing parameter

$$x = \frac{\Delta m}{\Gamma_\mathrm{B}} \approx 0.7 \tag{7.11.88}$$

where Δm is the mass difference of the heavy and light mass eigenstates for B_d^0–$\bar{\mathrm{B}}_\mathrm{d}^0$ system, and Γ_B is the average width of the two mass eigenstates. The observed process of ARGUS Group is

$$\mathrm{e}^+\mathrm{e}^- \to \Upsilon(4S) \to \mathrm{B}_\mathrm{d}^0\bar{\mathrm{B}}_\mathrm{d}^0 \to \mu^\pm\mu^\pm\mathrm{X} \tag{7.11.89}$$

Obviously, if we calibrate one $\mathrm{B}_\mathrm{d}^0 \to l^+\nu_l\mathrm{X}$, the flying out particle in other side must be $\bar{\mathrm{B}}_\mathrm{d}^0$, that is, at the same time t the decay product can only be $\mathrm{B}_\mathrm{d}^0\bar{\mathrm{B}}_\mathrm{d}^0$, because the angular momentum of $\Upsilon(4S)$ is 1, but $\mathrm{B}_\mathrm{d}^0\mathrm{B}_\mathrm{d}^0$ or $\bar{\mathrm{B}}_\mathrm{d}^0\bar{\mathrm{B}}_\mathrm{d}^0$ are identical particles, their total angular momentum can only be $0, 2, \ldots$. So the angular momentum conservation requires that the final state at the same time can only be $\mathrm{B}_\mathrm{d}^0\bar{\mathrm{B}}_\mathrm{d}^0$. Therefore, at one side there is $\mathrm{B}_\mathrm{d}^0 \to l^+\nu_l\mathrm{X}$, at the other side there must be $\bar{B}^0_d \to l^- \bar{\nu}_l X$, so the final state can only be $l^+ l^- X$. If we want to have $l^+ l^+ X$, only $\bar{B}^0_d$ oscillates into B^0_d then decays into $l^+ \nu_l X$. In that case, we can have the same sign events of $l^+ l^+ X$. Similarly we can have $l^- l^- X$ events. Note that when $\bar{B}^0_d$ transforms into B^0_d through mixing, B^0_d in other side already decayed, the final state cannot have $B^0_d B^0_d$ at the same time. So the existence of the same sign di-leptons verifies the existence of mixing. We need to point out here that in experiments the background caused by the same sign lepton events of the secondary decays must be ruled out. For example

$$\begin{array}{l} B^0_d \bar{B}^0_d \to [l^+ X_1][D\, X_2] \\ \qquad\qquad\qquad\qquad\;\; \lfloor\!\!\longrightarrow [l^+ X_3] \end{array} \tag{7.11.90}$$

l^+ here is from the decay of the charm meson D, it can be ruled out by kinematic method, for instance, by choosing events for which the l^+'s transverse momentum must be larger than some value.

ARGUS Group gives the mixing parameter $x \sim 0.7$ which is out of our expectation. People all thought before that $B^0_d - \bar{B}^0_d$ mixing should be very small, because they thought that the top quark mass $m_t < 30\,\text{GeV}$, therefore the small top quark mass makes the prediction of the standard model $x < 0.1$. The large mixing means large top quark mass. This issue firstly discussed by two group theorists [38]. They predicted that in order to explain the large mixing of B^0–$\bar{B}^0$, m_t must be lager than 50 GeV. This prediction means that people must go to higher energy accelerator to search for the top quark, this is very important conclusion. For searching for top quark Japanese National Laboratory (KEK) built e^+e^- collider of 60 GeV of the center-of-mass energy. They thought $m_t < 30\,\text{GeV}$, so this collider could produce $t\bar{t}$ pair. From this, we can see that $m_t < 30\,\text{GeV}$ was widely accepted by particle physics community. Ref. [37] played important role for pushing people to search for top quark at higher energy area.

According to Sec. 7.11.1, for B^0–$\bar{B}^0$ system, we can write down the following equations (see Eqs. (7.11.21) and (7.11.25))

$$\begin{cases} |B_H\rangle = p\,|B^0\rangle - q\,|\bar{B}^0\rangle \\ |B_L\rangle = p\,|B^0\rangle + q\,|\bar{B}^0\rangle \end{cases} \tag{7.11.91}$$

$$\begin{cases} |B^0(t)\rangle = g_+(t)\,|B^0\rangle + \dfrac{q}{p} g_-(t)\,|\bar{B}^0\rangle \\ |\bar{B}^0(t)\rangle = \dfrac{p}{q} g_-(t)\,|B^0\rangle + g_+(t)\,|\bar{B}^0\rangle \end{cases} \tag{7.11.92}$$

$$\begin{cases} \Delta m = m_{\rm H} - m_{\rm L} \\ \Delta\Gamma = \Gamma_{\rm H} - \Gamma_{\rm L} \\ \Gamma = \dfrac{1}{2}\left(\Gamma_{\rm H} + \Gamma_{\rm L}\right) \\ m = \dfrac{1}{2}\left(m_{\rm H} + m_{\rm L}\right) \\ x = \dfrac{\Delta m}{\Gamma} \end{cases} \tag{7.11.93}$$

where the expression of $g_\pm(t)$ can be found in Eq. (7.11.26). In what follows, without special notes we shall neglect the subscript d, and take $\mathrm{B}^0 \equiv \mathrm{B}^0_{\rm d}$, etc.

In experiments, the $\mathrm{B}^0\bar{\mathrm{B}}^0$ produced from the decay of $J^{PC} = 1^{--}$ $\Upsilon(4S)$ is a coherent state. Its state vector can be expressed according to its C parity± as

$$|\mathrm{B}^0\bar{\mathrm{B}}^0\rangle_{C=\pm} = \frac{1}{\sqrt{2}}\{|\mathrm{B}^0(\vec{k}, t_1)\bar{\mathrm{B}}^0(-\vec{k}, t_2)\rangle \pm |\mathrm{B}^0(-\vec{k}, t_2)\bar{\mathrm{B}}^0(\vec{k}, t_1)\rangle\} \tag{7.11.94}$$

Because the mass difference of the B^0 mass eigenstates $\mathrm{B}^0_{\rm H}$, $\mathrm{B}^0_{\rm L}$ Δm is very small, their width difference $\Delta\Gamma$ is also very small, when discussing the mixing of B^0–$\bar{\mathrm{B}}^0$, taking $\Delta\Gamma = 0$ is a very good approximation. Equation (7.11.94) becomes

$$\begin{aligned} \left|\mathrm{B}^0\bar{\mathrm{B}}^0\right\rangle_{C=\pm} \approx \frac{1}{\sqrt{2}} \mathrm{e}^{\mathrm{i}m(t_1+t_2) - \frac{\Gamma}{2}(t_1+t_2)} \Bigg\{ & \mathrm{i}\sin\frac{\Delta m}{2}(t_2 \pm t_1) \\ & \left[\frac{p}{q}|\mathrm{B}^0_{\vec{k}}\mathrm{B}^0_{-\vec{k}}\rangle \pm \frac{q}{p}|\bar{\mathrm{B}}^0_{\vec{k}}\bar{\mathrm{B}}^0_{-\vec{k}}\rangle\right] \\ & + \cos\frac{\Delta m}{2}(t_2 \pm t_1)\left[|\mathrm{B}^0_{\vec{k}}\bar{\mathrm{B}}^0_{-\vec{k}}\rangle \pm |\mathrm{B}^0_{-\vec{k}}\bar{\mathrm{B}}^0_{\vec{k}}\rangle\right]\Bigg\} \end{aligned} \tag{7.11.95}$$

From Eq. (7.11.95) we see that for $C = -1$ state, when $t_2 = t_1 = t$, as $\sin\frac{\Delta m}{2}(t - t) = 0$, so we only have $\mathrm{B}^0\bar{\mathrm{B}}^0$ and no $\mathrm{B}^0\mathrm{B}^0$, $\bar{\mathrm{B}}^0\bar{\mathrm{B}}^0$, that is what we discussed before.

Using Eq. (7.11.95) we can compute out $\langle l^\pm l^\pm \mathrm{X}|\mathrm{B}^0\bar{\mathrm{B}}^0\rangle_{C=\mp}$ and $\langle l^\pm l^\mp \mathrm{X}|\mathrm{B}^0\bar{\mathrm{B}}^0\rangle_{C=\mp}$. Under the assumption of

$$|\langle l^+l^+\mathrm{X}|\mathrm{B}^0_{\vec{k}}\mathrm{B}^0_{-\vec{k}}\rangle| = |\langle l^-l^-\mathrm{X}|\bar{\mathrm{B}}^0_{\vec{k}}\bar{\mathrm{B}}^0_{-\vec{k}}\rangle| = |\langle l^\pm l^\mp\mathrm{X}|\mathrm{B}^0_{\vec{k}}\bar{\mathrm{B}}^0_{-\vec{k}}\rangle| \tag{7.11.96}$$

we can calculate out the following decay width:

$$\begin{cases} \Gamma(\left(\mathrm{B}^0\bar{\mathrm{B}}^0\right)_{C=\mp} \to l^+l^+\mathrm{X}) \propto \mathrm{e}^{-\Gamma(t_1+t_2)} \left|\dfrac{p}{q}\right|^2 \sin^2 \dfrac{\Delta m}{2}(t_1 \mp t_2) \\ \Gamma(\left(\mathrm{B}^0\bar{\mathrm{B}}^0\right)_{C=\mp} \to l^-l^-\mathrm{X}) \propto \mathrm{e}^{-\Gamma(t_1+t_2)} \left|\dfrac{q}{p}\right|^2 \sin^2 \dfrac{\Delta m}{2}(t_1 \mp t_2) \\ \Gamma(\left(\mathrm{B}^0\bar{\mathrm{B}}^0\right)_{C=\mp} \to l^\pm l^\mp) \propto \mathrm{e}^{-\Gamma(t_1+t_2)} \cos^2 \dfrac{\Delta m}{2}(t_1 \mp t_2) \end{cases} \tag{7.11.97}$$

where in Eq. (7.11.97) we have neglected the phase space and the $\left|\frac{2}{\sqrt{2}}\right|^2$ factors, etc. Integrating Eq. (7.11.97) with respect to t_1, t_2 we have

$$\begin{cases} N_{++} \equiv N(\left[\mathrm{B}^0\bar{\mathrm{B}}^0\right]_{C=\mp} \to l^+l^+\mathrm{X}) \propto \left|\dfrac{p}{q}\right|^2 \left(1 - \dfrac{1 \pm x^2}{(1+x^2)^2}\right) \\ N_{--} \equiv N(\left[\mathrm{B}^0\bar{\mathrm{B}}^0\right]_{C=\mp} \to l^-l^-\mathrm{X}) \propto \left|\dfrac{q}{p}\right|^2 \left(1 - \dfrac{1 \pm x^2}{(1+x^2)^2}\right) \\ N_{+-} \equiv N(\left[\mathrm{B}^0\bar{\mathrm{B}}^0\right]_{C=\mp} \to l^\pm l^\mp\mathrm{X}) \propto \left(1 + \dfrac{1 \pm x^2}{(1+x^2)^2}\right) \end{cases} \tag{7.11.98}$$

Here we assume that $\Delta\Gamma \approx 0$, $\Delta\Gamma/\Gamma \ll 1$, and $x = \frac{\Delta m}{\Gamma}$.

In experiments, the quantity measured by ARGUS Group was

$$r = \left.\frac{N_{++} + N_{--}}{N_{+-}}\right|_{C=-} \approx \frac{x^2}{2+x^2} \tag{7.11.99}$$

In the denominator of r above N_{+-} should be understood as $N_{+-} + N_{-+}$, i.e. two times of the quantity in Eq. (7.11.98). Because in experiments we do not distinguish l^+l^- and l^-l^+ in the final states. At present, the newest measured value of x is [33]

$$x = 0.770 \pm 0.008 \tag{7.11.100}$$

7.11.3.2. *Theoretical calculation of the mixing parameter x*

Since B meson mass is approximately 5 GeV which is very heavy, so perturbation theory is applicable. The matrix elements M_{12} and Γ_{12} can be computed out from box diagrams (see Fig. 7.10). The definitions of M_{12}, Γ_{12} can be found in Sec. 7.11.1.

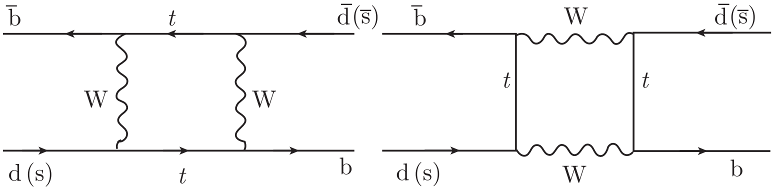


Fig. 7.10. Box diagrams of $\Delta B = 2$.

The result is [39]

$$M_{12} = \left(\frac{G_{\mathrm{F}}}{4\pi}\right)^2 \xi_{\mathrm{tq}}^2 M_{\mathrm{w}}^2 E\left(\frac{M_t^2(\mu)}{M_{\mathrm{w}}^2}\right) \eta_t^{\mathrm{B}}(\mu) \left\langle \mathrm{B}^0 \left|(\bar{\mathrm{q}}\mathrm{b})_{\mathrm{V-A}} (\bar{q}b)_{\mathrm{V-A}}\right| \bar{\mathrm{B}}^0 \right\rangle_{(\mu)} \tag{7.11.101}$$

where η_t^{B} is the calculated QCD radiative correction function at the energy scale μ. The function $E(x)$ is as follows:

$$\begin{aligned} E(x) &= \frac{4x - 11x^2 + x^2}{4(1-x)^2} - \frac{3x^3 \ln x}{4(1-x)^3} \\ x &= \frac{M_t^2(\mu)}{M_{\mathrm{w}}^2} \end{aligned} \tag{7.11.102}$$

$$\xi_{\mathrm{tq}} = V_{\mathrm{tb}} V_{\mathrm{tq}}^*, \quad \mathrm{q} = \mathrm{d} \text{ or } \mathrm{s} \tag{7.11.103}$$

In Eq. (7.11.101), the matrix element of the four quark operator is usually parameterized as the following form:

$$\left\langle \mathrm{B}^0 \right| (\bar{\mathrm{q}}\mathrm{b})_{\mathrm{V-A}} (\bar{\mathrm{q}}\mathrm{b})_{\mathrm{V-A}} \left|\bar{\mathrm{B}}^0\right\rangle_{(\mu)} = -\frac{4}{3}\mathrm{B}_{\mathrm{B}}(\mu) f_{\mathrm{B}}^2(\mu) m_{\mathrm{B}} \tag{7.11.104}$$

where f_{B} is the B meson decay constant, $(\bar{\mathrm{q}}\mathrm{b})_{\mathrm{V-A}} = \bar{\mathrm{q}}\gamma_\mu(1-\gamma_5)\mathrm{b}$, $\mathrm{B}_{\mathrm{B}}(\mu) = 1$ is called the vacuum saturation approximation, i.e. its value is obtained by putting vacuum state between the two $(\bar{\mathrm{q}}\mathrm{b})_{\mathrm{V-A}}$. Certainly, only inserting vacuum state is not enough, so $\mathrm{B}_{\mathrm{B}}(\mu) \neq 1$.

M_{12} should not depend on either the renormalization energy scale μ or on the renormalization scheme. So we redefine B_{B} and rewrite M_{12} as

$$\begin{aligned} M_{12} &= -\frac{G_{\mathrm{F}}^2 M_{\mathrm{w}}^2}{12\pi^2} \xi_{\mathrm{tq}}^2 E(x_{\mathrm{t}}) \eta_{\mathrm{t}}^{\mathrm{B}} f_{\mathrm{B}}^2 m_{\mathrm{B}} \mathrm{B}_{\mathrm{B}} \\ x_t &= \frac{M_t^2}{M_{\mathrm{w}}^2} \end{aligned} \tag{7.11.105}$$

It can be proved that [20]

$$\Delta M = 2\,|M_{12}| \tag{7.11.106}$$

Thus

$$x_d = \frac{(\Delta M)_{\mathrm{B_d}}}{\Gamma_{\mathrm{B_d}}} \approx \frac{2\,|M_{12}|}{\Gamma_{\mathrm{B_d}}} \approx 0.795\mathrm{B_B}\left(\frac{f_{\mathrm{B}}}{0.2\mathrm{GeV}}\right)^2 \tag{7.11.107}$$

The estimation of Eq. (7.11.107) is compatible with the experimental result.

Similarly, we can compute $x_{\mathrm{s}} = \frac{(\Delta M)_{\mathrm{B_s}}}{\Gamma_{\mathrm{B_s}}}$. The recent experimental result is [33]

$$x_{\mathrm{s}} = 26.94 \pm 0.29 \tag{7.11.108}$$

This means that $\mathrm{B_s^0}$–$\mathrm{\bar{B}_s^0}$ oscillation is very fast.

7.11.3.3. *B meson factory*

B meson factory means $\mathrm{e^+e^-}$ collider which can produce large amount of $\mathrm{B\bar{B}}$ pairs. For example, in order to produce large amount of $\mathrm{B_d^0\bar{B}_d^0}$ and $\mathrm{B_u^+B_u^-}$, the center-of-mass energy must concentrate on the resonate peak of $\Upsilon\,(4S)$ (its mass is 10.5794 GeV). $\Upsilon\,(4S)$ decays almost 100% into $\mathrm{B_d^0\bar{B}_d^0} + \mathrm{B_u^+B_u^-}$ (neutral and charged products are 50% of each).

The Cornell university electron–positron collider (CESR) and German electron positron collider (DESY) all can be used as B meson factory. Their $\mathrm{e^+e^-}$ beam energies are equal, so $\Upsilon\,(4S)$ produced at laboratory is almost at rest, therefore $\mathrm{B\bar{B}}$ from its decays are also almost at rest at laboratory. For this reason, we cannot measure the decay time of B and $\mathrm{\bar{B}}$ because their speed of flight is very small, so we cannot measure time-dependent physically observable quantifies.

In order to solve this problem, US and Japan built almost at the same time $\mathrm{e^+e^-}$ colliders — B factories of non-symmetrical beam energy. American Stanford University PEPII with beam energies 9 GeV $\mathrm{e^-}$ + 3.1 GeV $\mathrm{e^+}$, produced $\Upsilon\,(4S)$ flights with velocity $\beta\gamma = 0.56$; the distance between the two decay points of $\mathrm{B\bar{B}}$ is 260 μm. So we can measure the distance between the producing point of B or $\mathrm{\bar{B}}$ to the decay point of them by use of Silicon strip detector, afterwards we can extract the time of flight according to their velocity, so as to measure time-dependent physical quantities. This is the advantage of B meson factories with non-symmetrical beam energies. PEPII instant luminosity can reach $\mathcal{L} = 1.2 \times 10^{34}\,\mathrm{cm^{-2}\,s^{-1}}$.

For Japanese B factory (KEK-B), 8 GeV e^- + 3.5 GeV e^+, the velocity of the produced flying $\Upsilon(4S)$ $\beta\gamma = 0.43$. The distance between the two decay points of $B\bar{B}$ is 200 μm. The peak luminosity can reach $\mathcal{L} = 1.7 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$.

Each B factory can produce about $10^8 B\bar{B}$ every year. It can also produce large amount of secondary D meson, so they can do B physics research and also charm physics research. In recent years, the two B factories have obtained many physical achievements.

7.11.3.4. *CP violation in B meson system*

It is different from K^0–$\bar{K}^0$, for $B^0-\bar{B}^0$system the lifetime of the heavy meson B_H and that of the light meson B_L have not much difference, and their mass difference is also very small, $\Delta m = m_H - m_L = (3.337 \pm 0.033) \times 10^{-10}$ MeV (it is 100 times larger than Δm_K), so in experiments we cannot separate B_H beam from B_L one. Thus, it is no longer convenient to use $|\bar{\varepsilon}|$ for describing CP violation in B system.

In fact, the measurements of CP violation in B system are achieved by measuring partial decay rate asymmetry which is defined as

$$A_{CP}(t) = \frac{\Gamma(\bar{B}^0(t) \to \bar{f}) - \Gamma(B^0(t) \to f)}{\Gamma(\bar{B}^0(t) \to \bar{f}) + \Gamma(B^0(t) \to f)} \tag{7.11.109}$$

where $|\bar{f}\rangle = CP\,|f\rangle$, i.e. $|\bar{f}\rangle$ is the CP conjugate state of $|f\rangle$. When we cannot measure the decay time or we want to increase the statistics we can define the time-integrated CP asymmetry

$$A_{CP} = \frac{\int_0^\infty dt\Gamma(\bar{B}^0(t) \to \bar{f}) - \int_0^\infty dt\Gamma(B^0(t) \to f)}{\int_0^\infty dt\Gamma(\bar{B}^0(t) \to \bar{f}) + \int_0^\infty dt\Gamma(B^0(t) \to f)} \tag{7.11.110}$$

In what follows, we shall proceed discussions according to the CP violation classification given in Sec. 7.11.1.

(1) Direct CP violation(CP violation in decay amplitude): We first look at $B^\pm \to \pi^0 K^\pm$. This process involves charged $B^\pm$, B^+ B^- cannot mix, so CP violation can be only direct one, i.e. it can be defined as the difference of the absolute decay amplitudes squared,

$$A_{CP} = \frac{\left|A(B^+(t) \to \pi^0 K^+)\right|^2 - \left|A(B^-(t) \to \pi^0 K^-)\right|^2}{|A(B^+(t) \to \pi^0 K^+)|^2 + |A(B^-(t) \to \pi^0 K^-)|^2} \tag{7.11.111}$$

The preliminary result at present is [40]

$$A_{CP}(\mathrm{B}^+ \to \pi^0\mathrm{K}^+) = 0.050 \pm 0.025 \tag{7.11.112}$$

Obviously, this is not the final result, it needs further measurements, and the accuracy also needs improved.

(2) *Indirect CP violation* (*CP violation in mixing*): As an example we use the asymmetry of the same sign de-lepton final states to explain the characteristics of this kind of CP violation. In the standard model B^0 can only decay into l^+, $\bar{\mathrm{B}}^0$ can only decay into l^-, i.e. $\mathrm{B}^0 \to l^+\mathrm{X}$, $\bar{\mathrm{B}}^0 \to l^-\mathrm{X}$, so if there is no B^0–$\bar{\mathrm{B}}^0$ mixing, $\mathrm{B}^0\bar{\mathrm{B}}^0$ can only decay to opposite-sign di-lepton final states of l^+l^-, for instance,

$$\Upsilon(4S) \to \mathrm{B}^0\bar{\mathrm{B}}^0 \to l^+l^-\mathrm{X} \tag{7.11.113}$$

If there is B^0–$\bar{B}^0$ mixing, then the final states can have same sign di-lepton final states, such as, l^+l^+, l^-l^-, etc.:

$$\begin{tikzcd} \mathrm{B}^0 \arrow[dr] & \bar{\mathrm{B}}^0 \arrow[r] & \mathrm{B}^0 \to l^+\mathrm{X}, & \mathrm{B}^0 \arrow[dr] & \bar{\mathrm{B}}^0 \arrow[r] & l^-\mathrm{X} \\ & l^+\mathrm{X} & & & \bar{\mathrm{B}}^0 \to l^-\mathrm{X} & \end{tikzcd} \tag{7.11.114}$$

From Eq. (7.11.97) it is easy to obtain

$$\begin{aligned} A_{CP}(t) &= \frac{\Gamma\left((\mathrm{B}^0\bar{\mathrm{B}}^0)_{C=-} \to l^+l^+\mathrm{X}\right) - \Gamma\left((\mathrm{B}^0\bar{\mathrm{B}}^0)_{C=-} \to l^-l^-\mathrm{X}\right)}{\Gamma\left((\mathrm{B}^0\bar{\mathrm{B}}^0)_{C=-} \to l^+l^+\mathrm{X}\right) + \Gamma\left((\mathrm{B}^0\bar{\mathrm{B}}^0)_{C=-} \to l^-l^-\mathrm{X}\right)} \\ &= \frac{|p/q|^2 - |q/p|^2}{|p/q|^2 + |q/p|^2} = \frac{1 - |q/p|^4}{1 + |q/p|^4} \end{aligned} \tag{7.11.115}$$

Equation (7.11.115) shows that this kind of CP asymmetry is time independent. Obviously, according to Eq. (7.11.98) its time-integrated asymmetry is

$$A_{CP} = \frac{N_{++} - N_{--}}{N_{++} + N_{--}} = \frac{|p/q|^2 - |q/p|^2}{|p/q|^2 + |q/p|^2} = \frac{1 - |q/p|^4}{1 + |q/p|^4} \tag{7.11.116}$$

The result is the same as Eq. (7.11.115).

Because $|q/p|$ is not much different from 1, the value of Eq. (7.11.116) is very small, $A_{CP} \leqslant 10^{-4}$ [20]. In experiments, we have not measured out the same sign di-lepton CP asymmetry.

(3) *Mixed CP violation* (*CP violation in the interference of mixing and decay*): $B_d^0 \to \psi K_S$ is a typical example of mixed CP violation. This decay channel is the first and the most important process measured by B factories after their construction, and the measured results verified the prediction of the theories, i.e. in this channel CP asymmetry is larger.

When we discuss the CP violation in the decay of $B_d^0 \to \psi K_S$, we can neglect the CP violating effect in K^0–$\bar{K}^0$ system, because this effect is only the order of 10^{-3}, so we can take K_S, K_L as CP eigenstates, i.e. we can take the following approximation:

$$\begin{aligned} |K_S\rangle &\approx |K_1^0\rangle = \frac{1}{\sqrt{2}}\left(|K^0\rangle + |\bar{K}^0\rangle\right) \\ |K_L\rangle &\approx |K_2^0\rangle = \frac{1}{\sqrt{2}}\left(|K^0\rangle - |\bar{K}^0\rangle\right) \\ CP\,|K^0\rangle &= |\bar{K}^0\rangle \end{aligned} \tag{7.11.117}$$

Under this approximation

$$(|K_S\rangle)^+ = \langle K_S|\,, \quad (|K_L\rangle)^+ = \langle K_L| \tag{7.11.118}$$

The quark diagrams of $B_d^0 \to \psi K_S$ are shown in Fig. 7.11.

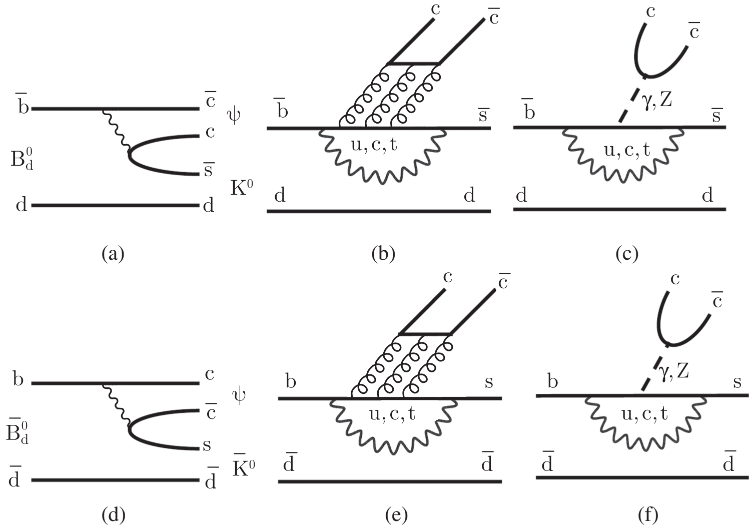


Fig. 7.11. $B_d^0, \bar{B}_d^0 \to \psi K_S$ quark diagrams: (a) and (d) are tree diagrams, (b) and (e) are QCD penguin diagrams, (c) and (f) are weak interaction penguin diagrams.

Now we calculate the CP asymmetry $A_{CP}(t)$. From Eq. (7.10.109), we obtain

$$\begin{aligned} A_{CP}(t) &= \frac{\Gamma\left(\bar{\mathrm{B}}^0(t)\to\bar{f}\right)-\Gamma\left(\mathrm{B}^0(t)\to f\right)}{\Gamma\left(\bar{\mathrm{B}}^0(t)\to\bar{f}\right)+\Gamma\left(\mathrm{B}^0(t)\to f\right)} \\ &= \frac{\left|A\left(\bar{\mathrm{B}}^0(t)\to\bar{f}\right)\right|^2-\left|A\left(\mathrm{B}^0(t)\to f\right)\right|^2}{\left|A\left(\bar{\mathrm{B}}^0(t)\to\bar{f}\right)\right|^2+\left|A\left(\mathrm{B}^0(t)\to f\right)\right|^2} \end{aligned} \tag{7.11.119}$$

where $A\left(\bar{\mathrm{B}}^0(t)\to\bar{f}\right)$ and $A\left(\mathrm{B}^0(t)\to f\right)$ are decay amplitudes. The final state is $|\mathrm{f}s\rangle=|\psi\mathrm{K}_\mathrm{S}\rangle$.

To simplify Eq. (7.11.119), define

$$\begin{aligned} &A(f)=\langle f|\,H_\mathrm{W}\left|\mathrm{B}^0\right\rangle,\bar{A}(f)=\langle f|\,H_\mathrm{W}\left|\bar{\mathrm{B}}^0\right\rangle \\ &A\left(\bar{f}\right)=\left\langle\bar{f}\right|H_\mathrm{W}\left|\mathrm{B}^0\right\rangle,\bar{A}\left(\bar{f}\right)=\left\langle\bar{f}\right|H_\mathrm{W}\left|\bar{\mathrm{B}}^0\right\rangle \\ &\lambda_f=\frac{q}{p}\frac{\bar{A}(f)}{A(f)},\bar{\lambda}_{\bar{f}}=\frac{p}{q}\frac{A\left(\bar{f}\right)}{\bar{A}\left(\bar{f}\right)} \end{aligned} \tag{7.11.120}$$

Using the expressions of $g_\pm(t)$ in Eqs. (7.11.91), (7.11.92) and (97.11.26), we can derive

$$\begin{aligned} &\left|A\left(\mathrm{B}^0(t)\to f\right)\right|^2=|A(f)|^2\,\frac{1}{2}\mathrm{e}^{-\Gamma t}\left(I_1+I_2\right) \\ &\left|A\left(\bar{\mathrm{B}}^0(t)\to f\right)\right|^2=\left|\frac{p}{q}\right|^2|A(f)|^2\,\frac{1}{2}\mathrm{e}^{-\Gamma t}\left(I_1-I_2\right) \\ &\left|A\left(\mathrm{B}^0(t)\to\bar{f}\right)\right|^2=\left|\frac{q}{p}\right|^2\left|\bar{A}\left(\bar{f}\right)\right|^2\frac{1}{2}\mathrm{e}^{-\Gamma t}\left(\bar{I}_1-\bar{I}_2\right) \\ &\left|A\left(\bar{\mathrm{B}}^0(t)\to\bar{f}\right)\right|^2=\left|\bar{A}\left(\bar{f}\right)\right|^2\frac{1}{2}\mathrm{e}^{-\Gamma t}\left(\bar{I}_1+\bar{I}_2\right) \end{aligned} \tag{7.11.121}$$

where

$$\begin{aligned} I_1&=\left(1+|\lambda_f|^2\right)\cosh\frac{\Delta\Gamma t}{2}+2\,\mathrm{Re}\left(\lambda_f\sinh\frac{\Delta\Gamma t}{2}\right) \\ I_2&=\left(1-|\lambda_f|^2\right)\cos\Delta mt-2\,\mathrm{Im}\left(\lambda_f\sin\Delta mt\right) \\ \bar{I}_1&=\left(1+\left|\bar{\lambda}_{\bar{f}}\right|^2\right)\cosh\frac{\Delta\Gamma t}{2}+2\,\mathrm{Re}\left(\bar{\lambda}_{\bar{f}}\sinh\frac{\Delta\Gamma t}{2}\right) \\ \bar{I}_2&=\left(1-\left|\bar{\lambda}\bar{f}\right|^2\right)\cos\Delta mt-2\,\mathrm{Im}\left(\bar{\lambda}_{\bar{f}}\sin\Delta mt\right) \end{aligned} \tag{7.11.122}$$

For the special final state $f=\psi\mathrm{K}_\mathrm{S}$, when neglecting the CP violating effect in K^0 meson, $\psi\mathrm{K}_\mathrm{S}$ will be CP eigenstate, and

$$\begin{aligned} &CP\,|\psi\mathrm{K}_\mathrm{S}\rangle=\eta_{CP}\,|\psi\mathrm{K}_\mathrm{S}\rangle=-\,|\psi\mathrm{K}_\mathrm{S}\rangle \\ &\text{i.e. }\eta_{CP}\left(\psi\mathrm{K}_\mathrm{S}\right)=-1 \end{aligned} \tag{7.11.123}$$

Therefore, for $f = \psi \mathrm{K_S}$, using Eqs. (7.11.121) and (7.11.122) we obtain

$$\begin{aligned} A_{CP}(t) &= \frac{\left|A\left(\bar{\mathrm{B}}^0(t) \to \psi \mathrm{K_S}\right)\right|^2 - \left|A\left(\mathrm{B}^0(t) \to \psi \mathrm{K_S}\right)\right|^2}{\left|A\left(\bar{\mathrm{B}}^0(t) \to \psi \mathrm{K_S}\right)\right|^2 + \left|A\left(\mathrm{B}^0(t) \to \psi \mathrm{K_S}\right)\right|^2} \\ &= -\frac{I_2}{I_1} = \frac{-(1 - |\lambda_f|^2)\cos \Delta mt + 2\,\mathrm{Im}\,(\lambda_f \sin \Delta mt)}{(1 + |\lambda_f|^2)\cosh \dfrac{\Delta\Gamma t}{2} + 2\,\mathrm{Re}(\lambda_f \sinh \dfrac{\Delta\Gamma t}{2})} \end{aligned} \tag{7.11.124}$$

For B^0–$\bar{\mathrm{B}}^0$ system, $\Delta\Gamma$ is very small, so for Eq. (7.11.124) taking $\Delta\Gamma \approx 0$ is a very good approximation. Thus Eq. (7.11.124) can be further simplified as

$$\begin{aligned} &A_{CP}(t) = S_f \sin \Delta mt - C_f \cos \Delta mt \\ &S_f = \frac{2\mathrm{Im}\lambda_f}{1 + |\lambda_f|^2}, C_f = \frac{1 - |\lambda_f|^2}{1 + |\lambda_f|^2} \end{aligned} \tag{7.11.125}$$

For $f = \psi \mathrm{K_S}$,

$$\lambda_f = \lambda_{\psi \mathrm{K_S}} = \frac{q}{p} \frac{A\left(\bar{\mathrm{B}}^0 \to \psi \mathrm{K_S}\right)}{A\left(\mathrm{B}^0 \to \psi \mathrm{K_S}\right)} \tag{7.11.126}$$

As long as we compute out $\lambda_{\psi \mathrm{K_S}}$, from Eq. (7.11.125) we can get *CP* asymmetry $A_{CP}(t)$.

We now calculate $\lambda_{\psi \mathrm{K_S}}$. It can be proved that [20], for B^0–$\bar{\mathrm{B}}^0$ system

$$|\Gamma_{12}| \ll |M_{12}| \tag{7.11.127}$$

Using Eqs. (7.11.127) and (7.11.38) we see that

$$\frac{q}{p} \approx \sqrt{\frac{M_{12}^*}{M_{12}}} = \frac{\xi_{\mathrm{td}}^*}{\xi_{\mathrm{td}}} = \frac{V_{\mathrm{tb}}^* V_{\mathrm{td}}}{V_{\mathrm{tb}} V_{\mathrm{td}}^*} = \mathrm{e}^{-2\mathrm{i}\beta} \tag{7.11.128}$$

In second step of Eq. (7.11.128) we have used the expression of M_{12} in Eq. (7.11.105) and

$$V_{\mathrm{td}} = |V_{\mathrm{td}}|\, \mathrm{e}^{-\mathrm{i}\beta} \tag{7.11.129}$$

Using the matrix expression of V_{CKM} in Eq. (7.7.6), we see that

$$\beta = \arg\left(-\frac{V_{\mathrm{cd}} V_{\mathrm{cb}}^*}{V_{\mathrm{td}} V_{\mathrm{tb}}^*}\right) = \arg\left(\frac{1}{V_{\mathrm{td}}}\right) \tag{7.11.130}$$

That is just the β angle of the unitarity triangle (see Sec. 7.11.3).

It can be proved that [20], if $A\left(\mathrm{B}^0 \to \psi \mathrm{K_S}\right)$ only involves one weak phase angle, then

$$\frac{\bar{A}\left(\bar{\mathrm{B}}_{\mathrm{d}}^0 \to \psi \mathrm{K_S}\right)}{A\left(\mathrm{B}_{\mathrm{d}}^0 \to \psi \mathrm{K_S}\right)} = \eta_{CP}\frac{\mathrm{e}^{-\mathrm{i}\varphi_{\mathrm{W}}}}{\mathrm{e}^{\mathrm{i}\varphi_{\mathrm{W}}}} = \eta_{CP}\mathrm{e}^{-2\mathrm{i}\varphi_{\mathrm{W}}} \tag{7.11.131}$$

where η_{CP} is the CP eigenvalue of $|\psi \mathrm{K_S}\rangle$, and $\eta_{CP} = -1$.

$$CP\left|\psi \mathrm{K_S}\right\rangle = \eta_{CP}\left|\psi \mathrm{K_S}\right\rangle \tag{7.11.132}$$

We now find the value of φ_{W} in Eq. (7.11.131).

From Fig. 7.11, as the contribution of the electroweak penguins (c) and (f) is very small so can be neglected, thus we only keep the contribution of the QCD penguin diagrams. We denote the contributions of the tree and QCD penguin diagrams as

$$\begin{aligned} T\left(\psi \mathrm{K_S}\right) &= V_{\mathrm{cb}}^* V_{\mathrm{cs}} t\left(\psi \mathrm{K_S}\right) \\ P_{\mathrm{u}}\left(\psi \mathrm{K_S}\right) &= V_{\mathrm{ub}}^* V_{\mathrm{us}} p_{\mathrm{u}}\left(\psi \mathrm{K_S}\right) \\ P_{\mathrm{c}}\left(\psi \mathrm{K_S}\right) &= V_{\mathrm{cb}}^* V_{\mathrm{cs}} p_{\mathrm{c}}\left(\psi \mathrm{K_S}\right) \\ P_{\mathrm{t}}\left(\psi \mathrm{K_S}\right) &= -V_{\mathrm{tb}}^* V_{\mathrm{ts}} p_{\mathrm{t}}\left(\psi \mathrm{K_S}\right) \end{aligned} \tag{7.11.133}$$

where we already factor out the CKM matrix elements. From the expression of V_{CKM} in Eq. (7.7.6) we can see that, up to λ^3 terms, $T\left(\psi \mathrm{K_S}\right)$, $P_{\mathrm{c}}\left(\psi \mathrm{K_S}\right)$, $P_{\mathrm{t}}\left(\psi \mathrm{K_S}\right)$ have same CKM factor(note that in Eq. (7.11.133) for the definition of P_{t} putting a minus sign is necessary), only $P_{\mathrm{u}}\left(\psi \mathrm{K_S}\right)$ has different CKM factor from P_{c}, P_{t}, T. But P_{u} term can be neglected because

$$\left|\frac{P_{\mathrm{u}}}{T}\right| \sim \lambda^2 \left|\frac{P_{\mathrm{u}}(\psi \mathrm{K_S})}{t(\psi \mathrm{K_S})}\right| \leqslant 1\% \tag{7.11.134}$$

So we get

$$A\left(\mathrm{B}_{\mathrm{d}}^0 \to \psi \mathrm{K_S}\right) \approx T\left(\psi \mathrm{K_S}\right) + P_{\mathrm{c}}\left(\psi \mathrm{K_S}\right) + P_{\mathrm{t}}\left(\psi \mathrm{K_S}\right) \tag{7.11.135}$$

On the right-hand side of Eq. (7.11.135) there is only one CKM factor which is real, so its phase $\varphi_{\mathrm{W}} = 0$, this leads to

$$\frac{\bar{A}\left(\bar{\mathrm{B}}_{\mathrm{d}}^0 \to \psi \mathrm{K_S}\right)}{A\left(\mathrm{B}_{\mathrm{d}}^0 \to \psi \mathrm{K_S}\right)} = \eta_{CP} = -1 \tag{7.11.136}$$

Substituting Eqs. (7.11.136) into (7.11.126) we get

$$\lambda_{\psi \mathrm{K_S}} = -\mathrm{e}^{-2\mathrm{i}\beta} \tag{7.11.137}$$

Substituting Eqs. (7.11.137) into (7.11.125) we have

$$A_{CP}(t) = \sin 2\beta \sin \Delta m t \tag{7.11.138}$$

As long as we measure the variation curve of $A_{CP}(t)$ with time t, we can get the value of $\sin 2\beta$. The recent world average number is [32]

$$\sin 2\beta = 0.679 \pm 0.020 \tag{7.11.139}$$

7.11.3.5. *Unitarity triangle*

Because V_{CKM} is a unitary matrix, its rows and columns are all complex orthogonal, i.e.

$$\sum_i V_{ik} V_{il}^* = \delta_{kl}, \quad \sum_i V_{ji} V_{ki}^* = \delta_{jk} \tag{7.11.140}$$

In detail, there are six orthogonal conditions, every orthogonal condition is a closed triangle in complex plain, this triangle is called unitarity triangle. It can be proved that the areas of all unitarity triangles are equal to each other and equal to $\frac{1}{2}J$, where J is called Jarlskog invariant quantity. Its definition is

$$\mathrm{Im}\left(V_{j\beta} V_{k\gamma} V_{j\gamma}^* V_{k\beta}^*\right) = J \sum_l \varepsilon_{jkl} \sum_\alpha \varepsilon_{\beta\gamma\alpha} \tag{7.11.141}$$

This quantity keeps invariant under the phase transformations of the matrix elements of V_{CKM}:

$$J = c_{12} c_{13}^2 c_{23} s_{12} s_{13} s_{23} \sin \delta_{13} \approx A^2 \lambda^6 \eta \tag{7.11.142}$$

where the definitions of c_{ij}, s_{ij}, δ_{ij}, A, λ, η can be found in Eqs. (7.7.5) and (7.7.6) for the expression of V_{CKM}. For detail, please see [20, Chapter 3].

J is related to the magnitude of CP violation. In principle, the side lengths and intersection angles of each unitarity triangle are relative to the experimentally observable quantities. But owing to the accuracy limitation of experiments, only these unitarity triangles whose three sides have comparable lengths can be easily measured. Only one unitarity triangle satisfies this condition, i.e.

$$V_{\mathrm{ud}} V_{\mathrm{ub}}^* + V_{\mathrm{cd}} V_{\mathrm{cb}}^* + V_{\mathrm{td}} V_{\mathrm{tb}}^* = 0 \tag{7.11.143}$$

This is a closed triangle in the complex plain as shown in Fig. 7.12.

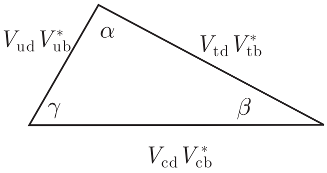


Fig. 7.12. Sketch of the unitarity triangle.

In the above, the tree angles α, β, γ are defined as

$$\begin{aligned}\alpha &= \arg\left(-\frac{V_{\rm td}V_{\rm tb}^*}{V_{\rm ud}V_{\rm ub}^*}\right) \equiv \phi_2\\ \beta &= \arg\left(-\frac{V_{\rm cd}V_{\rm cb}^*}{V_{\rm td}V_{\rm tb}^*}\right) \equiv \phi_1 \qquad (7.11.144)\\ \gamma &= \arg\left(-\frac{V_{\rm ud}V_{\rm ub}^*}{V_{\rm cd}V_{\rm cb}^*}\right) \equiv \phi_3\end{aligned}$$

For the measurements of the three angles, β is easy; it has been discussed in this section 7.11.3. The measurements of α and γ are a bit complicated. We shall not discuss them here. Readers who are interested in can refer [20]. Here we only give their 3σ world average values as [33]

$$\alpha = \left(91.0\begin{array}{l}+9.6\\-7.2\end{array}\right)^{\circ}, \quad \beta = \left(21.8^{\circ}\begin{array}{l}+2.8\\-1.9\end{array}\right)^{\circ}, \quad \gamma = \left(67.2^{\circ}\begin{array}{l}+6.9\\-9.1\end{array}\right)^{\circ}$$

7.11.4. D^0–$\bar{D}^0$ *mixing and CP violation*

At present, many laboratories in the world, for example, the e^+e^- colliders BEPCI, BEPCII, CESR-C, B factories PEP-II, KEK-B, etc., all produce large amount of charmed particles, the number of them is comparable with B_d–$\bar{B}_d$. Searching for D^0–$\bar{D}^0$ mixing and CP violation is not only possible but also important research topic.

7.11.4.1. *Theoretical prediction for* D^0–$\bar{D}^0$ *mixing*

According to Sec. 7.10.1, the evolution of the D^0–$\bar{D}^0$ system satisfies the following Schrödinger equation:

$$i\frac{d}{dt}\begin{pmatrix} D^0(t)\\ \bar{D}^0(t)\end{pmatrix} = \left(M - \frac{i}{2}\Gamma\right)\begin{pmatrix} D^0(t)\\ \bar{D}^0(t)\end{pmatrix} \qquad (7.11.145)$$

Assuming CPT invariance, we then can define

$$\begin{cases} |\mathrm{D}_1\rangle = p\left|\mathrm{D}^0\right\rangle + q\left|\bar{\mathrm{D}}^0\right\rangle \\ |\mathrm{D}_2\rangle = p\left|\mathrm{D}^0\right\rangle - q\left|\bar{\mathrm{D}}^0\right\rangle \\ CP\left|\mathrm{D}^0\right\rangle = \left|\bar{\mathrm{D}}^0\right\rangle \end{cases} \tag{7.11.146}$$

When CP conserves the CP parity of $|\mathrm{D}_1\rangle$ is positive, the CP parity of $|\mathrm{D}_2\rangle$ is negative. The eigenvalues of $|\mathrm{D}_1\rangle$ and $|\mathrm{D}_2\rangle$ are, respectively,

$$\lambda_1 = m_1 - \frac{\mathrm{i}}{2}\Gamma_1, \quad \lambda_2 = m_2 - \frac{\mathrm{i}}{2}\Gamma_2 \tag{7.11.147}$$

Put

$$\begin{aligned} &\Gamma = \frac{1}{2}(\Gamma_1 + \Gamma_2), m = \frac{1}{2}(m_1 + m_2) \\ &\Delta m = m_1 - m_2, \Delta\Gamma = \Gamma_1 - \Gamma_2 \\ &x = \Delta m/\Gamma, y = \Delta\Gamma/2\Gamma \end{aligned} \tag{7.11.148}$$

In the above equation, the definitions of Δm and $\Delta\Gamma$ have a minus sign difference from that of Eq. (7.11.23). This difference is convenient for comparison with experimental data. Theoretically, we can compute out the values of x and y.

In the standard model, the short-range contribution for D^0–$\bar{\mathrm{D}}^0$ mixing comes mainly from box diagram and dipenguin diagram. These are shown in Fig. 7.13 and Fig. 7.14.

In box diagrams, the part of exchanging b quark can be neglected owing to the suppression of very small value of V_{ub}, V_{cb}, so the contribution of the box diagrams comes mainly from exchanging s and d quarks. The detail calculation gives [41]

$$\Delta m_{\mathrm{D}}^{\mathrm{box}} \sim 0.5 \times 10^{-17}\left(\frac{m_{\mathrm{s}}\,(\mathrm{GeV})}{0.2\mathrm{GeV}}\right)^4 \left|\frac{f_{\mathrm{D}}}{f_\pi}\right|^2 \mathrm{GeV} \tag{7.11.149}$$

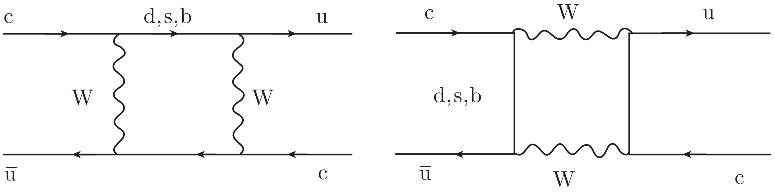


Fig. 7.13. Box diagrams for D^0–$\bar{\mathrm{D}}^0$ mixing.

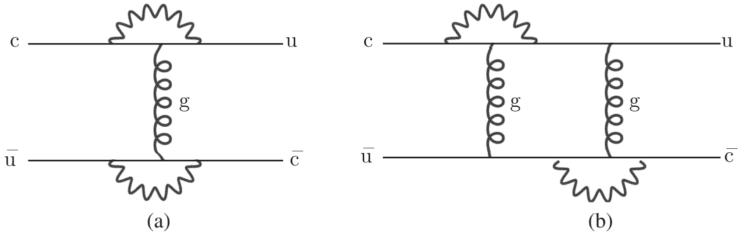


Fig. 7.14. Dipenguin diagrams: (a) short-range contribution; (b) long-range contribution from soft gluon momentum.

Using the total width of D^0 $\Gamma_{D^0} \sim 1.605 \times 10^{-12}$ GeV, from Eq. (7.11.149) we get

$$x^{\text{box}} = \frac{\Delta m_D^{\text{box}}}{\Gamma} \sim \mathcal{O}\left(10^{-7}\text{–}10^{-6}\right) \tag{7.11.150}$$

While the contribution of the dipenguin diagram (a) is [42]

$$\Delta m_D^{\text{dp}} \sim -0.2 \times 10^{-17} \left|f_D/f_\pi\right|^2 \tag{7.11.151}$$

It has the same order of magnitude but opposite sign comparing with box diagram Δm_D^{box}. On the whole, the short-range contribution

$$x^{\text{sd}} = \frac{\Delta m_D^{\text{sd}}}{\Gamma} \sim \mathcal{O}\left(10^{-7} - 10^{-6}\right) \tag{7.11.152}$$

As for the long-range contribution, because $m_c \sim 1.5$ GeV, the perturbation theory is not applicable, we do not have reliable computation method. The existed estimate in the literature gives [43]

$$x \sim y \sim 10^{-2} - 10^{-4} \tag{7.11.153}$$

7.11.4.2. *Experimental measurements for* D^0–$\bar{D}^0$ *mixing parameters*

At present, the experimental measurements have two kinds: one is to measure the time dependence of the final state decay width of D^0, $\bar{D}^0$ to KK, Kπ, $\pi\pi$, through fitting the measured data we can get x and y; the other is to measure the time-integrated total decay width to measure $x^2 + y^2$. The first kind needs to measure decay time, so it demands D^0 $\left(\bar{D}^0\right)$ to have large enough momentum. Usually, it needs momentum larger than 2.5 GeV/c. This kind of experiments can be done only at B meson factory through $D^* \to D\pi$ decay to produce high speed D^0. The second kind of experiments can be done at low energy e^+e^- colliders BEPC, CESR-C, here we shall not go to detail. The readers who are interested in can refer [20, Chapter 7]. Here we only give the world average value.

Heavy Flavor Averaging Group (HFAG) gives the world average values of the year 2011 (mixing parameter world average x and y) are [44]

$$\begin{aligned} x &= \left(0.63\begin{matrix}+0.19\\-0.26\end{matrix}\right) \times 10^{-2} \\ y &= (0.75 \pm 0.12) \times 10^{-2} \end{aligned} \tag{7.11.154}$$

7.11.4.3. *Direct CP violation in charmed particles*

From Eq. (7.11.154) we can see that D^0–$\bar{D}^0$ mixing is very small (the order of $\sim 10^{-3}$). So CP violations induced by mixing and by the interference of mixing and decay should be all very small. Up to now, the above two kinds of CP violations do not show any indications of existence. The only kind of a bit larger CP violation left is the direct CP violation. For example, $D^0 \to K^+K^-$, $D^0 \to \pi^+\pi^-$ can have direct CP violation which can be possibly observed.

The corresponding decay diagrams of $D^0 \to K^+K^-$ $D^0 \to \pi^+\pi^-$ are shown in Fig. 7.15.

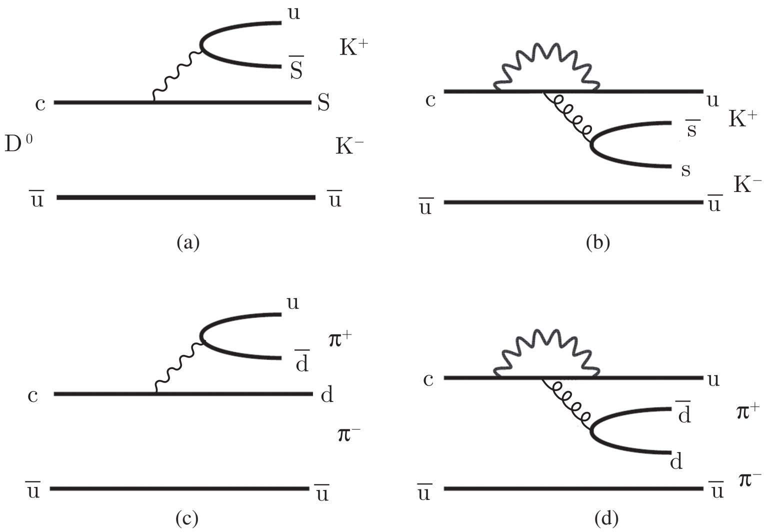


Fig. 7.15. $D^0 \to K^+K^-$, $\pi^+\pi^-$ decay diagrams.

The CP violation of $\mathrm{D}^0 \to \mathrm{f}$ can be divided to three parts:

$$\begin{aligned} A_{CP}(f) &= \frac{\Gamma\left(\mathrm{D}^0 \to \mathrm{f}\right) - \Gamma\left(\bar{\mathrm{D}}^0 \to \bar{\mathrm{f}}\right)}{\Gamma\left(\mathrm{D}^0 \to \mathrm{f}\right) + \Gamma\left(\bar{\mathrm{D}}^0 \to \bar{\mathrm{f}}\right)} \\ &= A_{CP}^{\mathrm{dir}}(\mathrm{f}) + A_{CP}^{\mathrm{m}}(\mathrm{f}) + A_{CP}^{\mathrm{i}}(\mathrm{f}) \end{aligned} \tag{7.11.155}$$

where $\mathrm{f} = \mathrm{K}^+\mathrm{K}^-$, $\pi^+\pi^-$ are CP eigenstates, the superscripts"dir" means direct CP violation, "m" means mixing induced CP violation, "i" means CP violation induced by the interference of mixing and decay. U spin analysis gives [45],

$$A_{CP}^{\mathrm{dir}}\left(\mathrm{D}^0 \to \mathrm{K}^+\mathrm{K}^-\right) = -A_{CP}^{\mathrm{dir}}\left(\mathrm{D}^0 \to \pi^+\pi^-\right) \tag{7.11.156}$$

$$A_{CP}^{\mathrm{m}}\left(\mathrm{D}^0 \to \mathrm{K}^+\mathrm{K}^-\right) = A_{CP}^{\mathrm{m}}\left(\mathrm{D}^0 \to \pi^+\pi^-\right) \tag{7.11.157}$$

$$A_{CP}^{\mathrm{i}}\left(\mathrm{D}^0 \to \mathrm{K}^+\mathrm{K}^-\right) = A_{CP}^{\mathrm{i}}\left(\mathrm{D}^0 \to \pi^+\pi^-\right) \tag{7.11.158}$$

So from Eqs. (7.11.155)–(7.11.158) we deduce

$$\Delta A_{CP} = A_{CP}\left(\mathrm{K}^+\mathrm{K}^-\right) - A_{CP}\left(\pi^+\pi^-\right) \approx 2A_{CP}^{\mathrm{dir}}\left(\mathrm{K}^+\mathrm{K}^-\right) \tag{7.11.159}$$

In the last step of deriving Eq. (7.11.159), the keypoint is the reliability of Eq. (7.11.156).

In 2012, LHCb Collaboration measured [46]

$$\Delta A_{CP} = (-0.82 \pm 0.21 \pm 0.11) \times 10^{-2} \tag{7.11.160}$$

where the first error is statistical, the second one is systematic, its signal significance is 3.5σ. Later, the result of LHCb is proved by CDF and Belle collaborations [47]. The average value of the above three groups is

$$\begin{aligned} \Delta A_{CP} &= A_{CP}(\mathrm{K}^+\mathrm{K}^-) - A_{CP}(\pi^+\pi^-) \\ &= (-0.74 \pm 0.15) \times 10^{-2} \end{aligned} \tag{7.11.161}$$

The main contribution to ΔA_{CP} comes from direct CP violation (see Eq. (7.10.159)). HFAG Group gives [48]

$$\Delta A_{CP}^{\mathrm{dir}} \approx (-0.678 \pm 0.147) \times 10^{-2} \tag{7.11.162}$$

its signal significance is 4.6σ. From Fig. 7.15, the tree diagrams are doubly Cabibbo suppressed, while the penguin diagrams are just singly Cabibbo suppressed. So singly Cabibbo suppressed process has larger direct CP violation. This is important for searching for new physics. But the measured result of LHCb Collaboration in 2014 gives [49],

$$\Delta A_{CP} = (+0.14 \pm 0.16(\text{stat.}) \pm 0.08(\text{syst.})) \times 10^{-2} \tag{7.11.163}$$

Comparing with Eq. (7.11.161) the two results have large difference, so we need more accurate measurements. Fortunately, in 2019 LHCb pronounced a new result [50]

$$\Delta A_{CP} = \{-18.2 \pm 3.2(\text{stat.}) \pm 0.9(\text{syst.})\} \times 10^{-4}$$

Combining with the previous result of this group they gave a new result which has signal significance larger than 5σ $\Delta A_{CP} = (-15.4 \pm 2.9) \times 10^{-4}$. This is the first result of CP violation observed in the world in charmed hadron decays, and that CP violating magnitude is compatible with the prediction of the standard model.

References

1. Pauli W. *Proc. Solvay Congr.*, 1933: 324.
2. Fermi E. *Z. Phys.*, 1934, 88: 161.
3. Gamov G, Teller E. *Phys. Rev.*, 1936, 49: 895.
4. Lee T D, Yang C N. *Phys. Rev.*, 1956, 104: 254.
5. Wu C S, *et al. Phys. Rev.*, 1957, 105: 1413; Garwin R L, Ledeman L M, Weinrich M. *Phys. Rev.*, 1957, 105: 1415; Telegdi V L, Friedman A M. *Phys. Rev.*, 1957, 105: 1681.
6. Feynman R P, Gell-Mann M. *Phys. Rev.*, 1958, 109: 193.
7. Cabibbo N. *Phys. Rev. Lett.*, 1963, 10: 531; *Phys. Rev. Lett.*, 1964, 12: 62.
8. Glashow S, Iliopoulos J, Maiani L. *Phys. Rev.*, 1970, D2: 1285.
9. Kobayashi M, Maskawa T. *Prog. Theor. Phys.*, 1973, 49: 652.
10. Christenson J H, *et al.*, *Phys. Rev. Lett.*, 1964, 13: 138.
11. Lee T D, Rosenbluth M, Yang C N. *Phys. Rev.*, 1949, 75: 9905.
12. Weinberg S. *Phys. Rev. Lett.*, 1967, 19: 1264; Glashow S L. *Nucl. Phys.*, 1961, 22: 579; Salam A. in Svartholm N, ed. *Elementary Particle Physics: Relativistic Groups and Analyticity*, Almquvist & Wiskell, Proceedings of the Eighth Nobel symposium, 1968: 367.
13. Nakamura K, *et al.* (Particle Data Group), *J. Phys. G: Nucl. Part. Phys.*, 2010, V37: 075021.
14. Thomas S A, Abdalla F B, Lahav O. *Phys. Rev. Lett.*, 2010, 105: 031301.

15. Xu G N, Yang Y T., *Nuclear Physics (the part of Nuclear Structure and Decay)*. Higher Education Press, 1989, (in Chinese).
16. Otten E W, Weinheimer C. *Rep. Prog. Phys.*, 2008, 71: 082601.
17. Hikasa K, *et al.* (Particle Data Group). *Phys. Rev.*, 1992, D45: S1.
18. Bai J Z, *et al.* (BES Collaboration). *Phys. Rev. Lett.*, 1992, 69: 3021.
19. Ting S C C, *et al. Phys. Rev. Lett.*, 1974, 33: 1404; Richter B, *et al. Phys. Rev. Lett.*, 1974, 33: 1406.
20. Du D S, *CP Violation*. Peking University Press, 2012 (in Chinese).
21. Wolfenstein L. *Phys. Rev. Lett.*, 1983, 51: 1945.
22. Gerstin S S, Zeldovich Y B. *Sov. Phys.*, 1956, JETP 2: 576.
23. Lee T D. *Particle Physics and Introduction to Field Theory*. Science Press, 1981: 614.
24. Sirlin A. *Rev. Mod. Phys.*, 1978, 50: 573.
25. McFarlane W K, *et al. Phys. Rev.*, 1985, D32: 547.
26. Thomas A W. arXiv:nucl-th/9609052.
27. Gell-Mann M, Levy M. *Nuovo Cimento*, 1960, 16: 705; Bernstein J, *et al. Nuovo Cimento*, 1960, 17: 757; Chou K C. *Sov. Phys.*, 1961, 12: 492 [JEPT, 1963, 39: 703].
28. Lee T D, Yang C N. *Phys. Rev.*, 1955, 98: 1501; Lee T D, Oehme R, Yang C N. *Phys. Rev.*, 1957, 106: 340.
29. Weisskopf V F, Wigner E P. *Z. Phys.*, 1930, 63: 54; 1930, 65: 18.
30. Kabir P K. *The CP Puzzle*. Academic Press, 1968, Appendix.
31. Gell-Mann M. *Phys. Rev.*, 1953, D92: 833.
32. Nakano T, Nishijima K. *Prog. Theor. Phys.*, 1953, 10: 580.
33. Beringer J, *et al.* (Particle Data Group). *Phys. Rev.*, 2012, D86: 010001.
34. Gewenger C, *et al. Phys. Lett.*, 1974, B48: 483; Fitch V L, *et al. Phys. Rev. Lett.*, 1973, 31: 1524; Williams H H, *et al. Phys. Rev. Lett.*, 1973, 31: 1521.
35. Wu T T, Yang C N. *Phys. Rev. Lett.*, 1964, 13: 380.
36. Herb J W, *et al. Phys. Rev. Lett.*, 1977, 39: 252.
37. Albrecht H, *et al.* (ARGUS Collab.) *Phys. Lett.*, 1987, B192: 245.
38. Du D S, Zhao Z Y. *Phys. Rev. Lett.*, 1987, 59: 1072; Ellis J, Hagelin J S, Rudaz S. *Phys. Lett.*, 1987, B192: 201.
39. Hagelin J S. *Nucl. Phys.*, 1981, B193: 123; Branco G C, Lavoura L, and Silva J P. *CP Violation*. Clarendon Press, 1999, Appendix B.
40. HFAG (Heavy Flavor Averaging Group) 2010.
41. Datta A, Kumbhakar D. *Z. Phys.*, 1985, C27: 515 ; Cheng H Y. *Phys. Rev.*, 1982, D26: 143; Donoghue J F, Golowich E, Volencia G. *Phys. Rev.*, 1986, D33: 1387.
42. Petrov A A. *Phys. Rev.*, 1997, D56: 1685.
43. Georgi H. *Phys. Lett.*, 1992, B297: 353; Bigi I I, Uraltsev N G. *Nucl. Phys.*, 2001, B592: 92; Donoghue J F, *et al. Phys. Rev.*, 1986, D33: 179.
44. Heavy Flavor Averaging Group, www.slac.stanford.edu/xorg/hfag/charm/March12/results_mixtcpv.html.
45. Grassman Y, Kagan A, Nir Y. *Phys. Rev.*, 2007, D75: 036008.
46. Aaij R, *et al.* (LHCb Collab.) *Phys. Rev. Lett.*, 2012, 108: 111602.

47. Aaltonen T, *et al.* (CDF Collab.) *Phys. Rev. Lett.*, 2012, 109: 111801; Ko B R, *et al.* (Belle Collab.), arXiv:1212.1975 (hep-ex).
48. Asner D, *et al.* (HFAG). arXiv:1010.1589 (hep-ex), www.slac.stanford.edu/xorg/hfag/charm/ICHEP12/DCPV/direct_indirect_cpv.html.
49. LHCb Collab., arXiv:1405.2797.
50. LHCb Collab. *Phys. ReV. Lett.*, 2019, 122: 211803.

This page intentionally left blank

Chapter 8

Gauge Theory of Electroweak Unification

Electromagnetic interaction is invariant under U(1) local gauge transformation. In Chapter 6, we obtained the quantum theory of electromagnetic interaction by extending the classical theory of electromagnetic interaction. Later, we clarified that the theory was U(1) local gauge invariant. Logically, we could also treat the problem in another way. If the Lagrangian $\mathcal{L}$ that describes the motion of electron is required to be local gauge invariant, there must be electromagnetic field. From this point of view, the electromagnetic field is a gauge field, and the electromagnetic force is a gauge force.

The symmetry group for electromagnetic interaction is U(1) group, which is Abelian group. In 1954, C.N. Yang and R.L. Mills extended the local gauge invariance from Abelian group to non-Abelian group — SU(2) group [1]. They discussed the possibility of keeping the gauge invariance for the interaction of p and n, a system composed of particles with isospin 1/2, and pointed out that, if the interaction is invariant under SU(2) local gauge transformation, then there must exist the relevant gauge field as a mediator for the interaction. The quanta of such gauge field may carry charges ±e and 0. This theory can be easily extended to SU(n) groups.

In 1960s, Glashow, Weinberg and Salam investigated to describe weak interaction with gauge field [2–4]. They unified electromagnetic and weak interactions with a larger symmetry group $\mathrm{SU(2)_L \times U(1)_Y}$. Weak interaction is also a gauge interaction like electromagnetic interaction. Electromagnetic and weak interaction are different aspects of a unified force — electromagnetic–weak force. The existence of the masses for particles may violate gauge symmetry in this model, so the mass terms in the original Lagrangian do not exist, i.e. the masses of all particles are zero. Particles obtain their masses through vacuum spontaneous symmetry breaking.

Next let us introduce the mechanism of the vacuum spontaneous symmetry breaking — the Higgs mechanism [5–10].

8.1. The Higgs Mechanism

Consider a system consisting of N real scalar fields ϕ_i, $i = 1, \ldots, N$. The Lagrangian of this system is

$$\mathcal{L}(x) = \sum_{j=1}^{N} \frac{1}{2}[\partial_\mu \phi_j][\partial^\mu \phi_j)] - U(\phi_i) \tag{8.1.1}$$

The first term in Eq. (8.1.1) is the kinetic energy term of the particles in the system, and the second term is the interaction potential. The scalar field ϕ_i is function of the space–time coordinate x. The potential is the function of the scalar fields.

Assume the minimum of the interaction potential of the system is not at $\phi_i = 0$. The lowest-energy state of an interacting system is usually defined as the vacuum state of the system. If the minimum value of the interaction potential of the system is not at the point $\phi_i = 0$, then the vacuum of the system does not coincide with the point $\phi_i = 0$. Suppose the minimum energy of the system occurs at $\phi_i = \phi_i^{(0)}$. Let us consider the case of a small vibration of the system near $\phi_i^{(0)}$. Expand the interaction potential $U(\phi_i)$ near $\phi_i^{(0)}$,

$$\begin{aligned} U(\phi_i) = U(\phi_i^{(0)}) &+ \frac{1}{2} \sum_{j,k=1}^{N} \left.\frac{\partial^2 U}{\partial \phi_j \partial \phi_k}\right|_{\phi_i = \phi_i^{(0)}} (\phi_j - \phi_j^{(0)})(\phi_k - \phi_k^{(0)}) \\ &+ \frac{1}{3!} \sum_{j,k,l=1}^{N} \left.\frac{\partial^3 U}{\partial \phi_j \partial \phi_k \partial \phi_l}\right|_{\phi_i = \phi_i^{(0)}} (\phi_j - \phi_j^{(0)}) \\ &\times (\phi_k - \phi_k^{(0)})(\phi_l - \phi_l^{(0)}) + \cdots \end{aligned} \tag{8.1.2}$$

Since $\phi_i^{(0)}$ is the point at which U takes the minimum value, then at this point

$$\left.\frac{\partial U}{\partial \phi_j}\right|_{\phi_i = \phi_i^{(0)}} = 0$$

so there is no term of the first power of $\left(\phi_j - \phi_j^{(0)}\right)$ in Eq. (8.1.2). Define

$$\left(M^2\right)_{jk} = \left.\frac{\partial^2 U}{\partial\phi_j \partial\phi_k}\right|_{\phi_i=\phi_i^{(0)}} \tag{8.1.3}$$

which is called the mass matrix of these N scalar fields. ϕ_j cannot be naively interpreted as fields of particles with definite masses, because there are non-diagonal elements in the mass matrix $\left(M^2\right)_{jk}$, i.e. the mass matrix is non-diagonal matrix. $\left(M^2\right)_{jk}$ is a real symmetric matrix, which can be diagonalized by unitary transformation. For a real matrix, the matrix of the unitary transformation required for diagonalizing it can be a real matrix. Let the unitary real matrix is R_{ij}. Then

$$\left(R^+\right)_{ij} = R_{ji} \tag{8.1.4}$$

Since R is the matrix that diagonalizes the mass matrix, so there is

$$\sum_{j,k=1}^{N} \left(R^+\right)_{ij} \left(M^2\right)_{jk} R_{kl} = M_i^2 \delta_{il} \tag{8.1.5}$$

Using the unitarity of R, we can solve $\left(M^2\right)_{jk}$ from Eq. (8.1.5)

$$\begin{aligned}\left(M^2\right)_{jk} &= \sum_{i,l=1}^{N} R_{ji} M_i^2 \delta_{il} \left(R^+\right)_{lk} \\ &= \sum_{i=1}^{N} R_{ji} M_i^2 \left(R^+\right)_{ik} \\ &= \sum_{i=1}^{N} M_i^2 R_{ji} R_{ki}\end{aligned} \tag{8.1.6}$$

Do the unitary transformation for $\phi_j - \phi_j^{(0)}$, and denote

$$\tilde{\phi}_i = \sum_{j=1}^{N} (\phi_j - \phi_j^{(0)}) R_{ji} \tag{8.1.7}$$

then the mass term can be changed as

$$\begin{aligned}
&\sum_{j,k=1}^{N} \left(M^2\right)_{jk} (\phi_j - \phi_j^{(0)})(\phi_k - \phi_k^{(0)}) \\
&\quad = \sum_{j,k=1}^{N} \sum_{i=1}^{N} M_i^2 R_{ji} R_{ki} (\phi_j - \phi_j^{(0)})(\phi_k - \phi_k^{(0)}) \\
&\quad = \sum_{i=1}^{N} M_i^2 \tilde{\phi}_i^2
\end{aligned} \tag{8.1.8}$$

Since ϕ_i^0 is the minimum point of $U(\phi_i)$, when ϕ_i changed to $\tilde{\phi}_i$, $\tilde{\phi}_i = 0$ will be the minimum point of $U(\tilde{\phi}_i)$. According to the property that the second derivative of an analytic function at the minimum point is not less than 0, we have

$$M_i^2 \geq 0 \tag{8.1.9}$$

Therefore, after the unitary transformation, the Lagrangian of the scalar field system becomes

$$\mathcal{L} = \sum_{i=1}^{N} \frac{1}{2} [\partial_\mu \tilde{\phi}_i \partial^\mu \tilde{\phi}_i - M_i^2 \tilde{\phi}_i^2] + O(\tilde{\phi}^3) - U(\phi_i^{(0)}) \tag{8.1.10}$$

The last term in Eq. (8.1.10) is a negligible constant term. The first term describes the scalar field with mass M_i, because such term can lead to Klein–Gordon equation

$$\partial_\mu \partial^\mu \tilde{\phi}_i + M_i^2 \tilde{\phi}_i = 0 \tag{8.1.11}$$

The cubic and higher-power terms of the scalar fields in Eq. (8.1.10) can be interpreted as the interaction terms for scalar particles.

Next we shall consider the case for a complex scalar field, and present a specific example for a scalar potential. The Lagrangian for the complex scalar field is

$$\mathcal{L} = \partial_\mu \phi^+ \partial^\mu \phi - U(\phi^+ \phi) \tag{8.1.12}$$

and the interaction potential is

$$U(\phi^+ \phi) = -\mu^2 \phi^+ \phi + \lambda(\phi^+ \phi)^2 \tag{8.1.13}$$

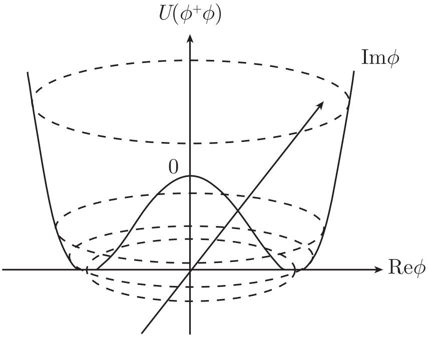


Fig. 8.1.

As both μ^2 and λ are real numbers larger than 0, the minimum value of the potential is not at the point $\phi = 0$. So $\phi = 0$ is not the vacuum state of the scalar system, as shown in Fig. 8.1.

To find the vacuum of the system, we only need to find the minimum value of the potential $U(\phi^+\phi)$. It is not difficult to find the minimum point

$$\phi_0 = \sqrt{\frac{\mu^2}{2\lambda}}\mathrm{e}^{\mathrm{i}\alpha} \equiv \frac{v}{\sqrt{2}}\mathrm{e}^{\mathrm{i}\alpha} \tag{8.1.14}$$

Without loss of generality, we can take $\alpha = 0$ for the arbitrary phase angle. Because there is U(1) invariance in the system, even if we take

$$\phi_0 = \frac{v}{\sqrt{2}}, \tag{8.1.15}$$

the arbitrary phase angle α can be recovered by U(1) transformation

$$\phi \to \mathrm{e}^{\mathrm{i}\alpha}\phi$$

The phase angle α is not an observable quantity in physics.

After the vacuum point of the system is fixed, the U(1) invariance of the system disappeared. The phenomenon that the symmetry of a system disappears due to the transformation of the physical vacuum point is called vacuum spontaneous symmetry breaking.

Considering the transformation of the physical vacuum point, we redefine the scalar field as

$$\phi(x) = \frac{1}{\sqrt{2}}[v + \chi(x) + \mathrm{i}\theta(x)] \tag{8.1.16}$$

Substituting Eq. (8.1.16) into Eq. (8.1.13), the interaction potential for scalar field, we can obtain

$$\begin{aligned}
U(\phi^+\phi) &= -\mu^2\phi^+\phi + \lambda(\phi^+\phi)^2 \\
&= -\frac{1}{2}\mu^2|\upsilon + \chi + \mathrm{i}\theta|^2 + \frac{1}{4}\lambda|\upsilon + \chi + \mathrm{i}\theta|^4 \\
&= -\frac{1}{2}\mu^2[(\upsilon + \chi)^2 + \theta^2] + \frac{1}{4}\lambda[(\upsilon + \chi)^2 + \theta^2]^2 \\
&= -\frac{1}{2}\mu^2[\upsilon^2 + 2\upsilon\chi + \chi^2 + \theta^2] + \frac{1}{4}\lambda[\upsilon^2 + 2\upsilon\chi + \chi^2 + \theta^2]^2 \\
&= \left(\frac{1}{4}\lambda\upsilon^4 - \frac{1}{2}\mu^2\upsilon^2\right) + \frac{1}{2}(\lambda\upsilon^2 - \mu^2)(2\upsilon\chi + \chi^2 + \theta^2) \\
&\quad + \lambda\upsilon^2\chi^2 + \lambda\upsilon\chi(\chi^2 + \theta^2) + \frac{1}{4}\lambda(\chi^2 + \theta^2)^2 \\
&= -\frac{1}{4}\upsilon^2\mu^2 + \lambda\upsilon^2\chi^2 + \lambda\upsilon\chi(\chi^2 + \theta^2) + \frac{1}{4}\lambda(\chi^2 + \theta^2)^2 \qquad (8.1.17)
\end{aligned}$$

We have used the relation

$$\upsilon = \sqrt{\mu^2/\lambda} \qquad (8.1.18)$$

defined in Eq. (8.1.14) in the last step of Eq. (8.1.17). Remove the constant term that does not contribute to the interaction, then the Lagrangian for the scalar field system can be expressed through the real scalar field χ and θ

$$\begin{aligned}
\mathcal{L} &= \frac{1}{2}\partial_\mu\chi\partial^\mu\chi + \frac{1}{2}\partial_\mu\theta\partial^\mu\theta - \lambda\upsilon^2\chi^2 - \lambda\upsilon\chi(\chi^2 + \theta^2) - \frac{1}{4}\lambda(\chi^2 + \theta^2)^2 \\
&= \frac{1}{2}\partial_\mu\chi\partial^\mu\chi - \frac{1}{2}(2\lambda\upsilon^2)\chi^2 + \frac{1}{2}\partial_\mu\theta\partial^\mu\theta - \lambda\upsilon\chi(\chi^2 + \theta^2) \\
&\quad - \frac{1}{4}\lambda(\chi^2 + \theta^2)^2 \qquad (8.1.19)
\end{aligned}$$

Equation (8.1.19) shows that the field χ obtained mass, but θ is still massless

$$m_\chi^2 = 2\lambda\upsilon^2, \quad m_\theta^2 = 0 \qquad (8.1.20)$$

After the symmetry breaking, the masses of the scalar fields χ and θ are

$$m_\chi^2 = \left.\frac{\partial^2 U}{\partial\chi^2}\right|_{\chi,\theta=0}, \quad m_\theta^2 = \left.\frac{\partial^2 U}{\partial\theta^2}\right|_{\chi,\theta=0} \qquad (8.1.21)$$

It can be verified by using Eq. (8.1.17) that Eq. (8.1.21) coincides with Eq. (8.1.20). It can be known from Eq. (8.1.16)

$$\begin{aligned}\mathrm{Re}\,\phi &= \frac{1}{\sqrt{2}}(v+\chi(x))\\ \mathrm{Im}\,\phi &= \frac{1}{\sqrt{2}}\theta(x)\end{aligned} \tag{8.1.22}$$

Therefore, it is shown from Fig. 8.1 that the second derivative of the interaction potential along the increasing direction of χ is not zero at the vacuum point after symmetry breaking. Physically, this is equivalent to χ having gained mass. But, along the increasing direction of θ, i.e. in the direction along the valley of the potential, the rotation symmetry still exists, and the second derivative of the potential is zero, so the scalar field θ is still massless. Here we can see that particle can gain mass after symmetry breaking. Simultaneously, massless particles may be generated in the direction perpendicular to the symmetry breaking. This is an inevitable result of spontaneous symmetry breaking. Massless particles generated in this way are referred to as Goldstone particle [11, 12].

8.2. Yang–Mills Theory

The quantum theory of the interaction of charged fermion and electromagnetic field can be set up by quantizing the classic electrodynamics. The theory obtained in this way is invariant under local U(1) gauge transformation. The physical meaning is, for a charged fermion and electromagnetic field interaction system, the phase angle of the fermion field can arbitrarily take different values at different coordinate point, the change caused by which can be compensated by the relevant change of the electromagnetic field, then the whole system is invariant. Therefore, the existence of electromagnetic field around fermions must be required theoretically as long as the interaction keeps invariant under local gauge transformation.

For a charged fermion field, the gauge transformation is actually the phase transformation, $\mathrm{e}^{ie\alpha(x)}$. Mathematically this transformation forms a U(1) group. The idea of gauge invariance can also be extended to the case of non-Abelian groups. Consider an isospin doublet composed of two fermion fields ψ_1 and ψ_2, which can be denoted as

$$\psi = \begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} \tag{8.2.1}$$

where ψ_1 and ψ_2 can be either proton and neutron, or u and d quarks, or neutrino and electron, etc. Without considering the fermion mass and the interaction between them, the Lagrangian density of this isospin doublet is

$$\mathcal{L} = \bar{\psi}\mathrm{i}\gamma^\mu \partial_\mu \psi \tag{8.2.2}$$

If substitute Eq. (8.2.1) into Eq. (8.2.2), this is just the Lagrangian for a system of two free fermion fields ψ_1 and ψ_2. This Lagrangian is invariant under SU(2) transformation

$$\psi \to \psi' = U\psi \equiv \mathrm{e}^{\mathrm{i}gT^i\beta^i}\psi \tag{8.2.3}$$

where g is the coupling constant, and β^i, $i = 1, 2, 3$, are arbitrary transformation parameters. Summation is indicated for repeated indices in Eq. (8.2.3). If no special instruction is given, all repeated indices mean summation. T^i's are the generators of SU(2) group, and $T^i = \frac{1}{2}\tau^i$. For basic representation, τ^i can be taken as Pauli matrix, $\tau^i = \sigma_i$. If β^i's are changed to be functions of space–time coordinate $\beta^i(x)$, then Eq. (8.2.2) is not invariant under local SU(2) transformation

$$\psi \to \psi' = U\psi \equiv \mathrm{e}^{\mathrm{i}gT^i\beta^i(x)}\psi \tag{8.2.4}$$

The reason is

$$\partial_\mu U \neq 0 \tag{8.2.5}$$

in this case. If requiring that the Lagrangian of the fermion doublet is invariant under SU(2) local gauge transformation of Eq. (8.2.4), then similar to the case of U(1) gauge transformation, gauge field A^i_μ relevant to SU(2) local gauge symmetry must be introduced. After gauge field A^i_μ introduced, the Lagrangian of Eq. (8.2.2) is extended to be

$$\mathcal{L} = \bar{\psi}\mathrm{i}\gamma^\mu(\partial_\mu - \mathrm{i}gT^iA^i_\mu)\psi \tag{8.2.6}$$

The gauge field in Eq. (8.2.6) should be transformed simultaneously as the fermion filed being under gauge transformation as Eq. (8.2.4), then the total Lagrangian is invariant. Let us derive the gauge transformation of the gauge field A^i_μ in the following. Let A'^i_μ be the gauge field after the gauge transformation. Then

$$\begin{aligned}\overline{\psi}'\mathrm{i}\gamma^\mu(\partial_\mu - \mathrm{i}gT^iA'^i_\mu)\psi' &= \overline{\psi}U^{-1}\mathrm{i}\gamma^\mu(\partial_\mu - \mathrm{i}gT^iA'^i_\mu)U\psi \\ &= \overline{\psi}\mathrm{i}\gamma^\mu U^{-1}(\partial_\mu - \mathrm{i}gT^iA'^i_\mu)U\psi \qquad (8.2.7)\end{aligned}$$

Equation (8.2.7) should be equal to the Lagrangian before gauge transformation, i.e.

$$\overline{\psi}\mathrm{i}\gamma^\mu U^{-1}(\partial_\mu - \mathrm{i}gT^iA'^i_\mu)U\psi = \overline{\psi}\mathrm{i}\gamma^\mu(\partial_\mu - \mathrm{i}gT^iA^i_\mu)\psi \qquad (8.2.8)$$

Comparing the left- and right-hand sides of Eq. (8.2.8), one can find the following equation for operator:

$$U^{-1}(\partial_\mu - \mathrm{i}gT^iA'^i_\mu)U = \partial_\mu - \mathrm{i}gT^iA^i_\mu \qquad (8.2.9)$$

Define covariant derivative as

$$\begin{aligned}D'_\mu &= \partial_\mu - \mathrm{i}gT^iA'^i_\mu \\ D_\mu &= \partial_\mu - \mathrm{i}gT^iA^i_\mu\end{aligned} \qquad (8.2.10)$$

then Eq. (8.2.9) can be concisely expressed as

$$D'_\mu = UD_\mu U^{-1} \qquad (8.2.11)$$

From Eq. (8.2.9), the transformation of the gauge field can be derived

$$T^iA'^i_\mu = UT^iA^i_\mu U^{-1} + \frac{\mathrm{i}}{g}U(\partial_\mu U^{-1}) \qquad (8.2.12)$$

There are kinetic terms of fermion fields and the interaction term of fermions and gauge field in the Lagrangian of Eq. (8.2.6), but without the mass terms of fermions. If there are mass terms, their general form should be like

$$\overline{\psi}M\psi \qquad (8.2.13)$$

where M is a 2×2 matrix. Such mass term usually violates gauge invariance, because the mass matrix generally cannot commute with the generators τ^i of SU(2) group, so

$$U^{-1}MU \neq M \qquad (8.2.14)$$

It is not difficult to see that the mass term of gauge field $A^i_\mu A^{i\mu}$ is not invariant either in the transformation of Eq. (8.2.12). Therefore, to keep gauge invariance in Yang–Mills theory, both fermion and gauge field must be massless. We will see later that, the left- and right-handed fields of fermion in the unified electroweak theory transform differently under gauge transformation, so the mass term of fermion must violate gauge invariance. Therefore, fermions must be massless in the original Lagrangian. The masses of fermion and gauge field can be generated through Higgs mechanism.

Similar to the case of the electromagnetic field, there should also be the kinetic term of the gauge field in a theory describing the interaction between the fermion doublet and gauge field. For the case of the electromagnetic field, the kinetic term is

$$-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{8.2.15}$$

where $F_{\mu\nu}$ is the electromagnetic field tensor,

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{8.2.16}$$

For non-Abelian gauge field, we can also find the kinetic energy term, and such kinetic energy term should be gauge invariant. We can still obtain the kinetic energy term of the non-Abelian gauge field by analogy with the electromagnetic field. For electromagnetic field, the covariant derivative is

$$D_\mu = \partial_\mu + \mathrm{i}eA_\mu \tag{8.2.17}$$

One can verify that the field tensor $F_{\mu\nu}$ and D_μ satisfy the relation

$$[D_\mu, D_\nu] = \mathrm{i}eF_{\mu\nu} \tag{8.2.18}$$

To find the the field tensor of non-Abelian gauge field, let us calculate the commutator of the covariant derivative of the non-Abelian gauge field in Eq. (8.2.10)

$$\begin{aligned}
[D_\mu, D_\nu] &= [\partial_\mu - \mathrm{i}gT^iA^i_\mu, \partial_\nu - \mathrm{i}gT^jA^j_\nu] \\
&= [\partial_\mu, -\mathrm{i}gT^jA^j_\nu] + [-\mathrm{i}gT^iA^i_\mu, \partial_\nu] + [-\mathrm{i}gT^iA^i_\mu, -\mathrm{i}gT^jA^j_\nu] \\
&= -\mathrm{i}gT^i(\partial_\mu A^i_\nu - \partial_\nu A^i_\mu) + (-\mathrm{i}g)^2[T^i, T^j]A^i_\mu A^j_\nu \\
&= -\mathrm{i}gT^i(\partial_\mu A^i_\nu - \partial_\nu A^i_\mu) - \mathrm{i}g^2\varepsilon^{ijk}A^i_\mu A^j_\nu T^k \\
&= -\mathrm{i}gT^i(\partial_\mu A^i_\nu - \partial_\nu A^i_\mu + g\varepsilon^{ijk}A^j_\mu A^k_\nu)
\end{aligned} \tag{8.2.19}$$

The commutation relation of the generators of SU(2) group has been used in Eq. (8.2.19)

$$[T^i, T^j] = \mathrm{i}\varepsilon^{ijk}T^k \tag{8.2.20}$$

It is known from Eq. (8.2.19) that, by analogy with Eq. (8.2.18), if define the relation between the non-Abelian gauge field tensor $F^i_{\mu\nu}$ and the covariant derivative as

$$[D_\mu, D_\nu] = -\mathrm{i}gF^i_{\mu\nu}T^i \tag{8.2.21}$$

then

$$F^i_{\mu\nu} = \partial_\mu A^i_\nu - \partial_\nu A^i_\mu + g\varepsilon^{ijk}A^j_\mu A^k_\nu \tag{8.2.22}$$

Next let us prove that the kinetic energy term of the non-Abelian gauge field can also be taken as

$$L_{\mathrm{NA}} = -\frac{1}{4}F^i_{\mu\nu}F^{i\mu\nu} \tag{8.2.23}$$

where summation is understood for the repeated indices, and such kinetic energy term is gauge invariant.[a] Using

$$\mathrm{Tr}[T^iT^j] = \frac{1}{2}\delta^{ij} \tag{8.2.24}$$

the kinetic energy term for gauge field can be written as

$$\mathcal{L}_{\mathrm{NA}} = -\frac{1}{2}\mathrm{Tr}\left[(F^i_{\mu\nu}T^i)(F^{j\mu\nu}T^j)\right] \tag{8.2.25}$$

With Eq. (8.2.11), the gauge transformation of covariant derivative, and Eq. (8.2.21), we can obtain the gauge transformation for the gauge field tensor

$$F'^i_{\mu\nu}T^i = UF^i_{\mu\nu}T^iU^{-1} \tag{8.2.26}$$

Then the gauge transformation of the kinetic energy term of gauge field is

$$\begin{aligned}\mathcal{L}'_{\mathrm{NA}} &= -\frac{1}{2}\mathrm{Tr}\left[(F'^i_{\mu\nu}T^i)(F'^{j\mu\nu}T^j)\right]\\ &= -\frac{1}{2}\mathrm{Tr}\left[UF^i_{\mu\nu}T^iU^{-1}UF'^{j\mu\nu}T^jU^{-1}\right]\\ &= -\frac{1}{2}\mathrm{Tr}\left[(F^i_{\mu\nu}T^i)(F^{j\mu\nu}T^j)\right]\\ &= \mathcal{L}_{\mathrm{NA}}\end{aligned} \tag{8.2.27}$$

That is, the kinetic energy term of gauge field, (8.2.23) is gauge invariant.

---

[a] The subscript in Eq. (8.2.23) "NA" denotes "non-Abelian".

Therefore, the Lagrangian that is invariant under SU(2) local gauge transformation is

$$\mathcal{L} = \overline{\psi}\mathrm{i}\gamma^\mu(\partial_\mu - \mathrm{i}gT^iA^i_\mu)\psi - \frac{1}{4}F^i_{\mu\nu}F^{i\mu\nu} \tag{8.2.28}$$

In order to know the properties of SU(2) gauge field further, let us look at the interaction term between the gauge field and fermion in Eq. (8.2.28)

$$g\bar{\psi}\gamma^\mu T^iA^i_\mu\psi$$

Substituting the explicit form of

$$T^1 = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad T^2 = \frac{1}{2}\begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad T^3 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

into this interaction term, we can get

$$\frac{1}{2}g\bar{\psi}\gamma^\mu\begin{pmatrix} A^3_\mu & A^1_\mu - \mathrm{i}A^2_\mu \\ A^1_\mu + \mathrm{i}A^2_\mu & -A^3_\mu \end{pmatrix}\psi = \frac{1}{2}g\bar{\psi}\gamma^\mu\begin{pmatrix} A^3_\mu & \sqrt{2}W^+ \\ \sqrt{2}W^- & -A^3_\mu \end{pmatrix}\psi \tag{8.2.29}$$

We have defined two new fields in Eq. (8.2.29),

$$W^\pm = \frac{A^1_\mu \mp \mathrm{i}A^2_\mu}{\sqrt{2}} \tag{8.2.30}$$

If what considered here is the doublet composed of neutrino and electron, i.e.

$$\psi = \begin{pmatrix} \nu \\ \mathrm{e} \end{pmatrix} \tag{8.2.31}$$

then Eq. (8.2.29) becomes

$$\frac{1}{2}g[\bar{\nu}\gamma^\mu A^3_\mu\nu + \sqrt{2}\bar{\nu}\gamma^\mu W^+_\mu\mathrm{e} + \sqrt{2}\bar{\mathrm{e}}\gamma^\mu W^-_\mu\nu - \bar{\mathrm{e}}\gamma^\mu A^3_\mu\mathrm{e}] \tag{8.2.32}$$

Every term in the Lagrangian in Eq. (8.2.32) should be electrically neutral. Therefore, from the couplings of $W^\pm_\mu$, A^3_μ, ν and e, one can judge that $W^\pm_\mu$ are electrically charged fields with charges of ±1, and A^3_μ is neutral. The quanta of these fields are usually denoted by $W^\pm$ and A^3. It should be noted that Eq. (8.2.32) can only be used as an example to judge the charges of $W^\pm$ and A^3, and it is not an exact form to describe the weak interaction involving electron and neutrinos. To express the weak interactions participated by electron and neutrino, ν and e in formula (8.2.32) should be replaced by the left-handed fields,

$$\begin{aligned} \nu_{\mathrm{L}} &= \frac{1}{2}(1-\gamma_5)\nu \\ \mathrm{e}_{\mathrm{L}} &= \frac{1}{2}(1-\gamma_5)\mathrm{e} \end{aligned} \tag{8.2.33}$$

which is known from (V–A) theory of weak interaction.

In conclusion, Yang–Mills theory shows that if the interaction is SU(2) local gauge invariant, there must be gauge fields to transfer the interaction, which is non-Abelian gauge fields. Three kinds of quanta for the gauge fields exist in Yang–Mills theory. Two kinds of them are charged particles with charges ±1, and the other is electrically neutral.

8.3. Glashow–Weinberg–Salam Electroweak Unification Model

Weak interaction and electromagnetic interaction are apparently different: at first, weak interaction is several orders weaker than electromagnetic interaction; secondly, the intermediate boson that transmits weak interaction is much heavier than photon that transmits electromagnetic interaction, so the force range of weak interaction is much shorter than that of electromagnetic interaction; at last, the intermediate bosons $\mathrm{W}^{\pm}$ are charged particles, while photon is electrically neutral. Although these two kinds of interactions are so different apparently, they can be described in a unified way by gauge theory and the idea of spontaneous symmetry breaking.

The particles that participate weak and electromagnetic interactions include leptons and quarks. According to the structure of the charged current with which these particles participating in weak interaction, these particles can be composed into a set of doublets. Each doublet is called one generation. Totally three generations for leptons have been found in experiment so far, which are

$$\begin{pmatrix} \nu_{\mathrm{e}} \\ \mathrm{e} \end{pmatrix}, \quad \begin{pmatrix} \nu_{\mu} \\ \mu \end{pmatrix}, \quad \begin{pmatrix} \nu_{\tau} \\ \tau \end{pmatrix}$$

and three generations for quarks, which are

$$\begin{pmatrix} \mathrm{u} \\ \mathrm{d} \end{pmatrix}, \quad \begin{pmatrix} \mathrm{c} \\ \mathrm{s} \end{pmatrix}, \quad \begin{pmatrix} \mathrm{t} \\ \mathrm{b} \end{pmatrix}$$

The behaviors of the left- and right-handed fields of fermion in weak interaction are different. For any fermion field ψ, its left-handed field is defined as

$$\psi_{\mathrm{L}} = \frac{1-\gamma_5}{2}\psi$$

and the definition of the right-handed field is

$$\psi_{\mathrm{R}} = \frac{1+\gamma_5}{2}\psi$$

We have known from Chapter 7 that weak interaction is V–A current–current interaction. For two fermion fields ψ_1 and ψ_2, the weak current is

$$j_\mu = \bar{\psi}_1\gamma_\mu(1-\gamma_5)\psi_2 \tag{8.3.1}$$

Define left- and right-handed projection operators as

$$L = \frac{1-\gamma_5}{2}, \quad R = \frac{1+\gamma_5}{2} \tag{8.3.2}$$

We can verify that there are the following properties for L and R:

$$L^2 = L, \quad R^2 = R, \quad LR = RL = 0 \tag{8.3.3}$$

According to the properties of left- and right-handed projection operators, the weak current in Eq. (8.3.1) can be rewritten as

$$\begin{aligned} j_\mu &= 2\bar{\psi}_1\gamma_\mu L\psi_2 \\ &= 2\bar{\psi}_1\gamma_\mu LL\psi_2 \\ &= 2\bar{\psi}_1 R\gamma_\mu L\psi_2 \\ &= 2\bar{\psi}_{1\mathrm{L}}\gamma_\mu\psi_{2\mathrm{L}} \end{aligned} \tag{8.3.4}$$

where the anti-commutation of γ_μ and γ_5 has been used. Equation (8.3.4) shows that only the left-handed field actually participates in the weak interaction in the weak current of V–A. Therefore, we need to distinguish the left- and right-handed fields of quarks and leptons in weak interaction. The left-handed field forms a doublet, while the right-handed field forms a singlet, which can ensure that the right-handed field does not contribute to the charged weak current.

The lepton doublet is

$$l_{\mathrm{L}}^{i}=\begin{pmatrix}\nu_{i\mathrm{L}}\\ l_{i\mathrm{L}}\end{pmatrix}=\frac{1-\gamma_5}{2}\begin{pmatrix}\nu_i\\ l_i\end{pmatrix} \tag{8.3.5}$$

where $i=\mathrm{e},\mu,\tau$. $\nu_{i\mathrm{L}}$ and $l_{i\mathrm{L}}$ denote the left-handed fields of three generations of leptons and quarks, respectively. Note that the symbol l_{L}^{i} with superscript i stands for the left-handed doublet, and $l_{i\mathrm{L}}$ with subscript i stands for the left-handed field for lepton. Under the SU(2) transformation, the left-handed doublet transforms as

$$l_{\mathrm{L}}^{\prime i}=\mathrm{e}^{\mathrm{i}gT^{j}\beta^{j}(x)}l_{\mathrm{L}}^{i} \tag{8.3.6}$$

where g is a constant, and the product of g and $\beta^{j}(x)$ makes up the transformation parameter for SU(2) group.

The right-handed lepton field is[b]

$$l_{i\mathrm{R}}=\frac{1+\gamma_5}{2}l_i$$

Under the SU(2) transformation, the right-handed field is singlet, i.e.

$$T^{j}l_{i\mathrm{R}}=0,\quad i,j=1,2,3 \tag{8.3.7}$$

where T^{j} is the singlet representation for the generator of SU(2) group, which is actually 0.

The left-handed doublet of quarks are

$$q_{\mathrm{L}}^{i}=\begin{pmatrix}u_{i\mathrm{L}}\\ d'_{i\mathrm{L}}\end{pmatrix}=\frac{1-\gamma_5}{2}\begin{pmatrix}u_i\\ d'_i\end{pmatrix} \tag{8.3.8}$$

where $i=1,2,3$. u_i stands for quarks u, c, t, respectively, and d'_i stands for the weak eigenstates of d, s, b. The reason that the down-type quarks in the left-handed doublets are chosen to be the weak eigenstates is that, according to Cabibbo theory, the mass eigenstates of quarks are different from their weak eigenstates, and the fields in the original Lagrangian should be weak eigenstates.

---

[b] Here we assume that right-handed neutrino does not exist in nature. Later we can see that this assumption will lead to massless neutrino. But it is not difficult to introduce the right-handed neutrinos and to gain the mini masses for neutrinos.

The right-handed fields for quarks are

$$u_{i\mathrm{R}} = \frac{1+\gamma_5}{2} u_i, \quad d'_{i\mathrm{R}} = \frac{1+\gamma_5}{2} d'_i$$

Under SU(2) transformation, the right-handed fields for quarks are singlets, and similar to Eq. (8.3.7) they satisfy

$$T^j u_{i\mathrm{R}} = 0, \quad T^j d'_{i\mathrm{R}} = 0 \tag{8.3.9}$$

The doublet and singlet defined above are called weak isospin doublet and singlet, respectively, which are different from isospin for strong interaction.

We still hypothesize that the charges and the third component of weak isospin for lepton and quarks satisfy Gell-Mann–Nishijima relation

$$Q = I_3 + \frac{Y}{2} \tag{8.3.10}$$

The quantum number Y in Eq. (8.3.10) is the weak hypercharge, and I_3 the third component of weak isospin. The isospin of the weak isospin doublet is 1/2, and the third components of the up and down components are ±1/2, respectively. The weak isospin of singlet is 0. We can derive the weak hypercharges for the doublet and singlet from the known electric charge and the third component of weak isospin of quarks and leptons. The weak hypercharge of lepton doublet is

$$Y(l^i_{\mathrm{L}}) = -1 \tag{8.3.11}$$

The weak hypercharge of lepton singlet is

$$Y(l_{i\mathrm{R}}) = -2 \tag{8.3.12}$$

The weak hypercharges of the quark left-handed doublet and right-handed singlet are

$$Y(q^i_{\mathrm{L}}) = \frac{1}{3}, \quad Y(u_{i\mathrm{R}}) = \frac{4}{3}, \quad Y(d_{i\mathrm{R}}) = -\frac{2}{3} \tag{8.3.13}$$

respectively.

Hypothesize that weak interaction is invariant under the local gauge transformation of $\mathrm{SU}(2)_{\mathrm{L}} \times \mathrm{U}(1)_{\mathrm{Y}}$, where L stands for the left-handed, and

Y the weak hypercharge. Then according to Yang–Mills theory, there must be SU(2) and U(1) gauge fields exist. The Lagrangian with $\mathrm{SU(2)_L \times U(1)_Y}$ local gauge symmetry is

$$\begin{aligned}\mathcal{L}_{\mathrm{FG}} = {} & \bar{l}^i_{\mathrm{L}} \mathrm{i}\gamma^\mu \left(\partial_\mu - \mathrm{i}gT^j A^j_\mu - \mathrm{i}\frac{g'}{2}\hat{Y}B_\mu\right) l^i_{\mathrm{L}} \\ & + \bar{q}^i_{\mathrm{L}} \mathrm{i}\gamma^\mu \left(\partial_\mu - \mathrm{i}gT^j A^j_\mu - \mathrm{i}\frac{g'}{2}\hat{Y}B_\mu\right) q^i_{\mathrm{L}} \\ & + \bar{l}_{i\mathrm{R}} \mathrm{i}\gamma^\mu \left(\partial_\mu - \mathrm{i}\frac{g'}{2}\hat{Y}B_\mu\right) l_{i\mathrm{R}} \\ & + \bar{u}_{i\mathrm{R}} \mathrm{i}\gamma^\mu \left(\partial_\mu - \mathrm{i}\frac{g'}{2}\hat{Y}B_\mu\right) u_{i\mathrm{R}} \\ & + \bar{d}'_{i\mathrm{R}} \mathrm{i}\gamma^\mu \left(\partial_\mu - \mathrm{i}\frac{g'}{2}\hat{Y}B_\mu\right) d'_{i\mathrm{R}} \end{aligned} \tag{8.3.14}$$

where g and g' are the coupling constants for $\mathrm{SU(2)_L}$ and $\mathrm{U(1)_Y}$ gauge fields, respectively. Note that a factor 1/2 has been separated from g' in Eq. (8.3.14), which is due to the fact that there are all factor 1/2 in the generators T^j of SU(2) group. If a factor of 1/2 is also separated from g', the relative relationship between g and g' will be easier to express. $\hat{Y}$ is the operator of weak hypercharge, which will result in the value of the hypercharge of the field when it is acted on the relevant field.

The $\mathrm{U(1)_Y}$ gauge transformations of the fermion and gauge field are

$$\begin{aligned} & \psi(x) \to \psi'(x) = \mathrm{e}^{\mathrm{i}\frac{g'}{2}Y\alpha(x)}\psi(x) \\ & B_\mu \to B'_\mu(x) = B_\mu(x) + \partial_\mu \alpha(x) \end{aligned} \tag{8.3.15}$$

where $\psi(x)$ stands for the doublet and singlet of the quark and leptons.

The $\mathrm{SU(2)_L}$ gauge transformation is

$$\begin{aligned} & \psi(x) \to \psi'(x) = \mathrm{e}^{\mathrm{i}gT^j\beta^j(x)}\psi(x) \\ & T^j A^j_\mu \to T^j {A_\mu}'^j = UT^j A^j_\mu U^{-1} + \frac{\mathrm{i}}{g}U(\partial_\mu U^{-1}) \\ & U \equiv \mathrm{e}^{\mathrm{i}gT^j\beta^j(x)} \end{aligned} \tag{8.3.16}$$

According to the discussion of Sec. 8.2, the Lagrangian in Eq. (8.3.14) is invariant under the above local gauge transformation.

Next we shall use the mass eigenstates of all the vector fields to express the Lagrangian in Eq. (8.3.14). Substitute the basic representation of the SU(2) generators $T^j = \frac{1}{2}\tau^j$ into the covariant derivative

$$\begin{aligned} D_\mu &= \partial_\mu - \mathrm{i}g\frac{\tau^j}{2}A^j_\mu - \mathrm{i}\frac{g'}{2}YB_\mu \\ &= \partial_\mu - \frac{\mathrm{i}}{2}\begin{pmatrix} gA^3_\mu + Yg'B_\mu & g(A^1_\mu - \mathrm{i}A^2_\mu) \\ g(A^1_\mu + \mathrm{i}A^2_\mu) & -gA^3_\mu + Yg'B_\mu \end{pmatrix} \\ &= \partial_\mu - \frac{\mathrm{i}}{2}\begin{pmatrix} gA^3_\mu + Yg'B_\mu & \sqrt{2}gW^+_\mu \\ \sqrt{2}gW^-_\mu & -gA^3_\mu + Yg'B_\mu \end{pmatrix} \end{aligned} \tag{8.3.17}$$

where the definition of the charged field $W^\pm_\mu$ can be seen in Eq. (8.2.30), and A^3_μ and B_μ are two neutral vector fields.

To keep the Lagrangian local gauge invariant, all the gauge fields are massless originally. But the intermediate boson $W^\pm_\mu$ that transmits weak interaction is actually massive. In Sec. 8.4 we will see that the gauge fields can obtain masses through vacuum spontaneous symmetry breaking. As the gauge fields obtain masses, one can see that both the neutral vector fields A^3_μ and B_μ are not their mass eigenstates. Their mass eigenstates are two orthogonal linear combinations of these neutral fields

$$\begin{aligned} A_\mu &= \cos\theta B_\mu + \sin\theta A^3_\mu \\ Z_\mu &= -\sin\theta B_\mu + \cos\theta A^3_\mu \end{aligned} \tag{8.3.18}$$

where θ is the Weinberg angle, which is defined in terms of g and $g\prime$,

$$\sin\theta = \frac{g'}{\sqrt{g^2 + g'^2}}, \quad \cos\theta = \frac{g}{\sqrt{g^2 + g'^2}} \tag{8.3.19}$$

The neutral fields B_μ and A^3_μ in Eq. (8.3.18) can be expressed in terms of mass eigenstates,

$$\begin{aligned} A^3_\mu &= \sin\theta A_\mu + \cos\theta Z_\mu \\ B_\mu &= \cos\theta A_\mu - \sin\theta Z_\mu \end{aligned} \tag{8.3.20}$$

Substituting the result of Eq. (8.3.20) into Eq. (8.3.17) and simplifying it, one can get

$$
\begin{aligned}
D_\mu &= \partial_\mu - \mathrm{i}g\frac{\tau^j}{2}A^j_\mu - \mathrm{i}\frac{g'}{2}YB_\mu \\
&= \partial_\mu - \frac{\mathrm{i}}{2}\begin{pmatrix}
\begin{array}{l}(g\sin\theta + Yg'\cos\theta)\,A_\mu \\ \quad + (g\cos\theta - Yg'\sin\theta)\,Z_\mu\end{array} & \sqrt{2}gW^+_\mu \\
\sqrt{2}gW^-_\mu & \begin{array}{l}(-g\sin\theta + Yg'\cos\theta)\,A_\mu \\ \quad - (g\cos\theta + Yg'\sin\theta)\,Z_\mu\end{array}
\end{pmatrix} \\
&= \partial_\mu - \frac{\mathrm{i}}{2}\begin{pmatrix}
\begin{array}{l}g\sin\theta\,(1+Y)\,A_\mu \\ \quad +g\dfrac{\cos^2\theta - Y\sin^2\theta}{\cos\theta}Z_\mu\end{array} & \sqrt{2}gW^+_\mu \\
\sqrt{2}gW^-_\mu & \begin{array}{l}g\sin\theta\,(-1+Y)\,A_\mu \\ \quad -g\dfrac{\cos^2\theta + Y\sin^2\theta}{\cos\theta}Z_\mu\end{array}
\end{pmatrix}
\end{aligned}
\tag{8.3.21}
$$

Considering Eqs. (8.3.20) and (8.3.21), and substituting the definitions of the left-handed doublet and right-handed field, and the hypercharges in Eqs. (8.3.11)–(8.3.13) into Eq. (8.3.21), Eq. (8.3.14) becomes

$$
\begin{aligned}
\mathcal{L}_{\mathrm{FG}} = {}& \bar{l}^i_{\mathrm{L}}\mathrm{i}\gamma^\mu\partial_\mu l^i_{\mathrm{L}} + \bar{l}^i_{\mathrm{L}}\gamma^\mu\begin{pmatrix}
\dfrac{g}{2\cos\theta}Z_\mu & \dfrac{g}{\sqrt{2}}W^+_\mu \\
\dfrac{g}{\sqrt{2}}W^-_\mu & -g\sin\theta A_\mu - g\dfrac{\cos 2\theta}{2\cos\theta}Z_\mu
\end{pmatrix} l^i_{\mathrm{L}} \\
& + \bar{q}^i_{\mathrm{L}}\mathrm{i}\gamma^\mu\partial_\mu q^i_{\mathrm{L}} + \bar{q}^i_{\mathrm{L}}\gamma^\mu \\
& \times\begin{pmatrix}
\begin{array}{l}\dfrac{2}{3}g\sin\theta A_\mu \\ \quad +g\dfrac{\cos^2\theta - \frac{1}{3}\sin^2\theta}{2\cos\theta}Z_\mu\end{array} & \dfrac{g}{\sqrt{2}}W^+_\mu \\
\dfrac{g}{\sqrt{2}}W^-_\mu & \begin{array}{l}-\dfrac{1}{3}g\sin\theta A_\mu \\ \quad -g\dfrac{\cos^2\theta + \frac{1}{3}\sin^2\theta}{2\cos\theta}Z_\mu\end{array}
\end{pmatrix} q^i_{\mathrm{L}} \\
& + \bar{l}_{i\mathrm{R}}\mathrm{i}\gamma^\mu\partial_\mu l_{i\mathrm{R}} + \bar{l}_{i\mathrm{R}}\gamma^\mu\left(-g\sin\theta A_\mu + g\frac{\sin^2\theta}{\cos\theta}Z_\mu\right) l_{i\mathrm{R}}
\end{aligned}
$$

$$
\begin{aligned}
&+\bar{u}_{i\mathrm{R}}\mathrm{i}\gamma^{\mu}\partial_{\mu}u_{i\mathrm{R}}+\bar{u}_{i\mathrm{R}}\gamma^{\mu}\left(\frac{2}{3}g\sin\theta A_{\mu}-\frac{2}{3}g\frac{\sin^{2}\theta}{\cos\theta}Z_{\mu}\right)u_{i\mathrm{R}}\\
&+\bar{d}'_{i\mathrm{R}}\mathrm{i}\gamma^{\mu}\partial_{\mu}\bar{d}'_{i\mathrm{R}}+\bar{d}'_{i\mathrm{R}}\left(-\frac{1}{3}g\sin\theta A_{\mu}+\frac{1}{3}g\frac{\sin^{2}\theta}{\cos\theta}Z_{\mu}\right)d'_{i\mathrm{R}}\\
=\;&\bar{\nu}_{i\mathrm{L}}\mathrm{i}\gamma^{\mu}\partial_{\mu}\nu_{i\mathrm{L}}+\bar{l}_{i}\mathrm{i}\gamma^{\mu}\partial_{\mu}l_{i}+\bar{u}_{i}\mathrm{i}\gamma^{\mu}\partial_{\mu}u_{i}+\bar{d}'_{i}\mathrm{i}\gamma^{\mu}\partial_{\mu}d'_{i}\\
&+\frac{g}{2\sqrt{2}}\left[\bar{l}_{i}\gamma^{\mu}\left(1-\gamma_{5}\right)\nu_{i}+\bar{d}'_{i}\gamma^{\mu}\left(1-\gamma_{5}\right)u_{i}\right]W_{\mu}^{-}\\
&+\frac{g}{2\sqrt{2}}\left[\bar{\nu}_{i}\gamma^{\mu}\left(1-\gamma_{5}\right)l_{i}+\bar{u}_{i}\gamma^{\mu}\left(1-\gamma_{5}\right)d'_{i}\right]W_{\mu}^{+}\\
&+\frac{g}{4\cos\theta}\Bigg[\bar{\nu}_{i}\gamma^{\mu}\left(1-\gamma_{5}\right)\nu_{i}+\bar{l}_{i}\gamma^{\mu}\left(-(1-4\sin^{2}\theta)+\gamma_{5}\right)l_{i}\\
&+\bar{u}_{i}\gamma^{\mu}\left(\left(1-\frac{8}{3}\sin^{2}\theta\right)-\gamma_{5}\right)u_{i}\\
&+\bar{d}'_{i}\gamma^{\mu}\left(-\left(1-\frac{4}{3}\sin^{2}\theta\right)+\gamma_{5}\right)d'_{i}\Bigg]Z_{\mu}\\
&+g\sin\theta\left[-\bar{l}_{i}\gamma^{\mu}l_{i}+\frac{2}{3}\bar{u}_{i}\gamma^{\mu}u_{i}-\frac{1}{3}\bar{d}'_{i}\gamma^{\mu}d'_{i}\right]A_{\mu}
\end{aligned}
\tag{8.3.22}
$$

The first four terms in Eq. (8.3.22) are the kinetic energy terms of the leptons and quarks. The fifth and sixth terms are charged weak current terms which are transmitted by the charged intermediate boson $\mathrm{W}^{\pm}$, and these two terms are Hermitian conjugates of each other. They describe the interaction between fermion and the charged intermediate boson. The seventh term is the neutral weak current term, which is transmitted by the neutral intermediate boson Z^{0}. This is a new weak interaction predicted by the non-Abelian local gauge theory. The non-Abelian gauge theory shows that weak interaction is not only induced by charged current, but also by neutral current, then the type of particles in weak interaction may not change either. Except that the weak neutral currents for neutrinos are in the form of V–A, the weak neutral current for quark and leptons are not pure V–A currents, the general form of which is $\gamma_{\mu}(g_{\mathrm{V}}-g_{\mathrm{A}}\gamma_{5})$, where g_{V} and g_{A} are all functions of Weinberg angle. In principle, g_{V}, g_{A} and then the Weinberg angle can be determined by neutrino–lepton scattering experiment

$$\nu_{\mathrm{e}}+\mathrm{e}\rightarrow\nu_{\mathrm{e}}+\mathrm{e}$$

The experimental measured value is

$$\sin^2\theta \approx 0.23 \tag{8.3.23}$$

The last term in Eq. (8.3.22) is in the form of vector current coupling, and later we shall see that the vector field A_μ is massless. All these properties are consistent with electromagnetic field, therefore this term is just the term of electromagnetic coupling, and the coupling constant of which is the electromagnetic coupling constant e, i.e.

$$e = g\sin\theta \tag{8.3.24}$$

The relation of the coupling constants g, g' and the electromagnetic coupling constant e can be depicted in Fig. 8.2.

Substituting Eq. (8.3.24) into Eq. (8.3.22), we can obtain the part of the Lagrangian for the interaction of fermion and gauge fields in terms of the weak and electromagnetic couplings

$$\begin{aligned}\mathcal{L}_{\mathrm{FG}} = {} & \bar{\nu}_{i\mathrm{L}}\mathrm{i}\gamma^\mu\partial_\mu\nu_{i\mathrm{L}} + \bar{l}_i\mathrm{i}\gamma^\mu\partial_\mu l_i + \bar{u}_i\mathrm{i}\gamma^\mu\partial_\mu u_i + \bar{d}'_i\mathrm{i}\gamma^\mu\partial_\mu d'_i \\ & + \frac{g}{2\sqrt{2}}\left[\bar{l}_i\gamma^\mu\left(1-\gamma_5\right)\nu_i + \bar{d}'_i\gamma^\mu\left(1-\gamma_5\right)u_i\right]W_\mu^- \\ & + \frac{g}{2\sqrt{2}}\left[\bar{\nu}_i\gamma^\mu\left(1-\gamma_5\right)l_i + \bar{u}_i\gamma^\mu\left(1-\gamma_5\right)d'_i\right]W_\mu^+ \\ & + \frac{g}{4\cos\theta}\left[\bar{\nu}_i\gamma^\mu\left(1-\gamma_5\right)\nu_i + \bar{l}_i\gamma^\mu\left(-(1-4\sin^2\theta)+\gamma_5\right)l_i\right. \\ & + \bar{u}_i\gamma^\mu\left(\left(1-\frac{8}{3}\sin^2\theta\right)-\gamma_5\right)u_i \\ & \left. + \bar{d}'_i\gamma^\mu\left(-\left(1-\frac{4}{3}\sin^2\theta\right)+\gamma_5\right)d'_i\right]Z_\mu \\ & + e\left[-\bar{l}_i\gamma^\mu l_i + \frac{2}{3}\bar{u}_i\gamma^\mu u_i - \frac{1}{3}\bar{d}'_i\gamma^\mu d'_i\right]A_\mu\end{aligned} \tag{8.3.25}$$

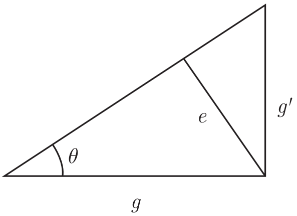


Fig. 8.2. The relationship of the coupling constants.

The last term in Eq. (8.3.25) is for the electromagnetic current of quarks and leptons, and the coupling of which with electromagnetic field is e. The coefficients of the sub-terms for each particle are -1, $\frac{2}{3}$, $-\frac{1}{3}$, which are the quantum numbers of the charges of lepton and quarks, respectively. This is not accidental result, but an inevitable result of Gell-Mann–Nishijima relation. The coupling coefficient of the weak neutral current is not the same as that of weak charged current, which is different by a factor $\frac{1}{\sqrt{2}\cos\theta}$.

Equation (8.3.25) is the consequence of the theory of local gauge symmetry $\mathrm{SU(2)_L \times U(1)_Y}$. To keep the local gauge symmetry, fermions and gauge fields must all be massless, but in reality all particles except photon are all with masses.[c] In the standard model, gauge boson and fermion masses are obtained by the vacuum spontaneous symmetry breaking mechanism, the Higgs mechanism. The combination of the $\mathrm{SU(2)_L \times U(1)_Y}$ electroweak unification theory and QCD, which will be introduced in the next chapter, is called the standard model of particle physics.

8.4. The Vacuum Spontaneous Symmetry Breaking: The Higgs Mechanism

Hypothesize there is a kind of scalar field $\Phi(x)$, which interacts with fermions and gauge fields in a way respecting local gauge symmetry. By assuming an appropriate interaction potential for the scalar field, the state with the magnitude of the scalar field being zero is not the lowest-energy state of the system. In quantum theory, the vacuum of a system is the state with the lowest energy for the system. Therefore, the state with the magnitude of the field being zero is not the vacuum state of the system, and in the vacuum state of the system the magnitude of the scalar field is not zero. In the vacuum state of the scalar field system, the interaction of scalar field with gauge field and fermions does not physically appear as the scalar field participating in the interaction. In this case, the scalar field system is in the "vacuum" state, and the existence of the quantum of the scalar field cannot be detected physically. Therefore, the effect of this interaction is only manifest itself as the non-zero mass obtained by the gauge field and fermions.

Let us analyze in detail the mechanism in the electroweak unification model that enables the gauge field particles and fermions to obtain their masses.

---

[c] There is now experimental evidence that neutrinos have tiny masses.

In order to construct $\mathrm{SU(2)_L \times U(1)_Y}$ local gauge invariant interaction between scalar field $\Phi(x)$ and gauge fields by using covariant derivative $D_\mu = \partial_\mu - \mathrm{i}gT^jA^j_\mu - \mathrm{i}\frac{g'}{2}YB_\mu$, the scalar field $\Phi(x)$ must be in a form of doublet, since the generators T^j's of gauge group are 2×2 matrices in its basic representative. Later we shall see that the simplest case for the hypercharge of the scalar doublet that can construct $\mathrm{SU(2)_L \times U(1)_Y}$ gauge invariant interaction between the scalar doublet and fermions is

$$Y = 1 \tag{8.4.1}$$

So we shall discuss the case $Y = 1$ for the scalar doublet. We still hypothesize that the Gell-Mann–Nishijima relation still holds for scalar field, i.e. the charge of scalar field satisfies

$$Q = I_3 + \frac{Y}{2} \tag{8.4.2}$$

Then we can obtain that the charge of the scalar field with the third component of weak isospin $I_3 = \frac{1}{2}$ is $Q = 1$, and the charge $Q = 0$ for the third component $I_3 = -\frac{1}{2}$, therefore we can explicitly write the scalar doublet as

$$\Phi(x) = \begin{pmatrix} \phi^+(x) \\ \phi^0(x) \end{pmatrix} \tag{8.4.3}$$

where both $\phi^+(x)$ and $\phi^0(x)$ are complex fields. The real degree of freedom is 4 for the scalar doublet $\Phi(x)$, and we will use $(\phi^+)^+ = \phi^-$ later.

The Lagrangian that is $\mathrm{SU(2)_L \times U(1)_Y}$ local gauge invariant is

$$(D_\mu\Phi)^+ D^\mu\Phi$$

To realize the vacuum spontaneous symmetry breaking, a potential is needed

$$V(\Phi) = -\mu^2\Phi^+\Phi + \lambda(\Phi^+\Phi)^2 \tag{8.4.4}$$

where both μ and λ are real numbers that are larger than 0. The minimum value of the potential is not at $\Phi = 0$. So the Lagrangian for the interaction between the scalar and gauge field is

$$\mathcal{L}_H = (D_\mu\Phi)^+ D^\mu\Phi - V(\Phi) \tag{8.4.5}$$

here Φ is defined as the eigenvalue of the field operator $\hat{\Phi}$ at the eigenstate of the scalar field $|\Phi\rangle$.

$$\hat{\Phi}\,|\Phi\rangle = \Phi\,|\Phi\rangle \tag{8.4.6}$$

therefore Φ is a complex number. According to Eq. (8.4.4), the minimum of $V(\Phi)$ is at the point

$$\Phi^{+}\Phi|_{\min} = \frac{\mu^2}{2\lambda} \tag{8.4.7}$$

that is

$$\Phi|_{\min} = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ \upsilon \end{pmatrix}, \quad \upsilon \equiv \sqrt{\mu^2/\lambda} \tag{8.4.8}$$

The state of the scalar field at this point is defined as the vacuum state, i.e.

$$\Phi|_{\min} = \langle 0|\,\hat{\Phi}\,|0\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ \upsilon \end{pmatrix} \tag{8.4.9}$$

Equation (8.4.9) is equivalent to the condition that only Φ^0 has non-zero vacuum expectation value

$$\langle 0|\,\hat{\Phi}^0\,|0\rangle = \frac{\upsilon}{\sqrt{2}}$$

Only such choice can maintain the electromagnetic U(1) symmetry after the vacuum spontaneous symmetry breaking, because the charge of Φ^0 is zero.

Several fields with zero vacuum expectation value can be defined according to Eq. (8.4.9)

$$\begin{aligned}\Phi(x) &\equiv \begin{pmatrix} \phi^{+}(x) \\ [\upsilon + H(x) + \mathrm{i}\chi(x)]/\sqrt{2} \end{pmatrix} \\ &= \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ \upsilon \end{pmatrix} + \begin{pmatrix} \phi^{+}(x) \\ [H(x) + \mathrm{i}\chi(x)]/\sqrt{2} \end{pmatrix} \end{aligned} \tag{8.4.10}$$

where H and χ are all real scalar field with zero vacuum expectation value. If substitute Eq. (8.4.10) into $\mathcal{L}_H$, then the $\mathrm{SU(2)_L} \times \mathrm{U(1)_Y}$ symmetry of the Lagrangian $\mathcal{L}_H$ will be violated because of the existence of the zero vacuum expectation value of the scalar field

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ \upsilon \end{pmatrix}$$

but the electromagnetic U(1) symmetry still exists. For example, the term $\Phi^+\Phi$ is invariant under the joint transformation of $\mathrm{SU}(2)_\mathrm{L}$

$$\Phi(x) \to \Phi'(x) = U_2\Phi(x) \equiv \mathrm{e}^{\mathrm{i}gT^j\beta^j(x)}\Phi(x) \tag{8.4.11}$$

and $\mathrm{U}(1)_\mathrm{Y}$

$$\Phi(x) \to \Phi'(x) = U_1\Phi(x) \equiv \mathrm{e}^{\mathrm{i}\frac{g'}{2}Y\alpha(x)}\Phi(x) \tag{8.4.12}$$

for $\Phi(x)$, but after the new definition of Eq. (8.4.10)

$$\begin{aligned}\Phi^+\Phi = {} & \begin{pmatrix}\phi^+\\ [H+\mathrm{i}\chi]/\sqrt{2}\end{pmatrix}^+\begin{pmatrix}\phi^+\\ [H+\mathrm{i}\chi]/\sqrt{2}\end{pmatrix} + \begin{pmatrix}\phi^+\\ [H+\mathrm{i}\chi]/\sqrt{2}\end{pmatrix}^+\frac{1}{\sqrt{2}}\begin{pmatrix}0\\ \upsilon\end{pmatrix}\\ & + \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ \upsilon\end{pmatrix}^+\begin{pmatrix}\phi^+\\ [H+\mathrm{i}\chi]/\sqrt{2}\end{pmatrix} + \frac{1}{2}\upsilon^2\end{aligned} \tag{8.4.13}$$

Apparently, the second and third term are not invariant under the transformation of

$$\begin{pmatrix}\phi^+\\ [H+\mathrm{i}\chi]/\sqrt{2}\end{pmatrix} \to U_1U_2\begin{pmatrix}\phi^+\\ [H+\mathrm{i}\chi]/\sqrt{2}\end{pmatrix} \tag{8.4.14}$$

therefore the $\mathrm{SU}(2)_\mathrm{L}\times\mathrm{U}(1)_\mathrm{Y}$ symmetry of the term $\Phi^+\Phi$ disappears. Expand Eq. (8.4.13) further

$$\Phi^+\Phi = \phi^+\phi^- + \frac{1}{2}\left(H^2+\chi^2\right) + \upsilon H + \frac{1}{2}\upsilon^2 \tag{8.4.15}$$

The $\mathrm{U}(1)_\mathrm{em}$ electromagnetic transformation of ϕ^+, ϕ^-, H and χ in Eq. (8.4.15) is

$$\phi^+ \to \mathrm{e}^{\mathrm{i}e\hat{Q}\alpha(x)}\phi^+ = \mathrm{e}^{\mathrm{i}e\alpha(x)}\phi^+ \tag{8.4.16a}$$

$$\phi^- \to \mathrm{e}^{\mathrm{i}e\hat{Q}\alpha(x)}\phi^- = \mathrm{e}^{-\mathrm{i}e\alpha(x)}\phi^- \tag{8.4.16b}$$

$$H \to \mathrm{e}^{\mathrm{i}e\hat{Q}\alpha(x)}H = H \tag{8.4.16c}$$

$$\chi \to \mathrm{e}^{\mathrm{i}e\hat{Q}\alpha(x)}\chi = \chi \tag{8.4.16d}$$

where $\hat{Q}$ in the above equations are the charge operator. The fields H and χ keep invariant under $\mathrm{U}(1)_\mathrm{em}$ transformation, because H and χ are neutral fields, so $\hat{Q}H = 0$, $\hat{Q}\chi = 0$. It is not difficult to see that Eq. (8.4.15) is invariant under $\mathrm{U}(1)_\mathrm{em}$ transformation of Eqs. (8.4.16a)–(8.4.16d). Similar to this, if we substitute Eq. (8.4.10) into $\mathcal{L}_H$, then we can find that the symmetry is broken from $\mathrm{SU(2)_L \times U(1)_Y}$ down to $\mathrm{U(1)_{em}}$ symmetry. U(1) symmetry in electromagnetic interaction is exactly what the real world holds.

Substituting the scalar field after vacuum spontaneous symmetry breaking $\Phi(x)$ in Eq. (8.4.10) into $\mathcal{L}_H$ of Eq. (8.4.5), we can obtain the interactions between scalar and gauge fields.

8.4.1. *The kinetic energy, mass and self-interaction term of scalar field*

Neglecting the gauge fields in the covariant derivative, then the Lagrangian for pure scalar field interaction can be obtained

$$\mathcal{L}_H = (\partial_\mu \Phi)^+ \partial^\mu \Phi - V(\Phi) \tag{8.4.17}$$

Let us treat the term of scalar potential at first. Substituting Eq. (8.4.15) into Eq. (8.4.4) and simplifying it, then we can get

$$\begin{aligned} V(\Phi) = {} & \mu^2 H^2 + 2v\lambda \left(\phi^+\phi^- + \frac{1}{2}H^2 + \frac{1}{2}\chi^2\right) H \\ & + \lambda \left(\phi^+\phi^- + \frac{1}{2}H^2 + \frac{1}{2}\chi^2\right)^2 + \frac{1}{4}\lambda v^4 - \frac{1}{2}\mu^2 v^2 \end{aligned} \tag{8.4.18}$$

and the kinetic term is

$$\begin{aligned} (\partial_\mu \Phi)^+ \partial^\mu \Phi &= \left(\partial_\mu \phi^-, \frac{1}{\sqrt{2}}\partial_\mu H - \frac{\mathrm{i}}{\sqrt{2}}\partial_\mu \chi\right) \begin{pmatrix} \partial^\mu \phi^+ \\ \frac{1}{\sqrt{2}}\partial^\mu H + \frac{\mathrm{i}}{\sqrt{2}}\partial^\mu \chi \end{pmatrix} \\ &= \partial_\mu \phi^- \partial^\mu \phi^+ + \frac{1}{2}\partial_\mu H \partial^\mu H + \frac{1}{2}\partial_\mu \chi \partial^\mu \chi \end{aligned} \tag{8.4.19}$$

Combine these two terms, then the part for the pure scalar fields can be obtained

$$\begin{aligned} \mathcal{L}_H = {} & \partial_\mu \phi^- \partial^\mu \phi^+ + \frac{1}{2}\partial_\mu H \partial^\mu H + \frac{1}{2}\partial_\mu \chi \partial^\mu \chi - \mu^2 H^2 \\ & - 2v\lambda \left(\phi^+\phi^- + \frac{1}{2}H^2 + \frac{1}{2}\chi^2\right) H - \lambda \left(\phi^+\phi^- + \frac{1}{2}H^2 + \frac{1}{2}\chi^2\right)^2 \end{aligned} \tag{8.4.20}$$

The unimportant constant term has been neglected in the above equation. Equation (8.4.20) shows that the scalar field H has gained mass, and $\phi^\pm$

and χ are still massless. $\phi^{\pm}$ and χ are Goldstone particles. The kinetic and mass terms of the neutral scalar field H should be in the following form:

$$\frac{1}{2}\partial_\mu H \partial^\mu H - \frac{1}{2}m_H^2 H^2$$

Comparing the above form and the fourth term of Eq. (8.4.20), we can obtain

$$m_H = \sqrt{2}\mu \tag{8.4.21}$$

There are totally four scalar particles — H, ϕ^+, ϕ^-, χ in $\mathcal{L}_H$, where only H gained mass, meanwhile $\phi^{\pm}$ and χ are Goldstone particles with zero mass. Are they all observable particles in physics?

The degree of freedom of the scalar doublet $\Phi(x)$ is 4. However Eq. (8.4.10) is not the only way to define the vacuum expectation value for the scalar field. Different from Eq. (8.4.10), $\Phi(x)$ can always be expressed as

$$\Phi(x) = \mathrm{e}^{-\mathrm{i}gT^j\phi^j(x)} \begin{pmatrix} 0 \\ [v + H(x)]/\sqrt{2} \end{pmatrix} \tag{8.4.22}$$

mathematically, where ϕ^j, $j = 1, 2, 3$, and $H(x)$ are all real fields with zero vacuum expectation values. Their total number is four, which is consistent with the degree of freedom. Since the Lagrangian is invariant under $\mathrm{SU(2)_L}$ transformation, we can do an SU(2) gauge transformation on the scalar field doublet with the transformation parameter taken as $\beta^j(x) = \phi^j(x)$, then

$$\begin{aligned} \Phi(x) &\to U_2\Phi(x) \\ &= \mathrm{e}^{\mathrm{i}gT^j\phi^j(x)}\Phi(x) \\ &= \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \end{aligned} \tag{8.4.23}$$

Gauge transformation does not change the physics of the system, because the system is gauge invariant. However, compared with Eq. (8.4.10), there are no scalar fields $\phi^{\pm}$ and χ in Eq. (8.4.23), which suggests that $\phi^{\pm}$ and χ are just gauge degrees of freedom, and they are not the physical particles. Only the quantum of the scalar field $H(x)$ is physical particle. $H(x)$ is referred to as Higgs field, and its quantum is Higgs particle.

Equation (8.4.23) is equivalent to selecting a specific gauge, where only physical particles are included. Thus the number of fields to be treated in the theory is greatly reduced. This gauge is referred to as unitary gauge. We shall continue our discussion in unitary gauge in the following.

In the unitary gauge, substituting

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \upsilon + H(x) \end{pmatrix} \tag{8.4.24}$$

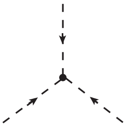
into Eq. (8.4.17) again, which is equivalent to removing the terms involving the scalar fields $\phi^{\pm}$ and χ in Eq. (8.4.20), then one can get

$$\begin{aligned} \mathcal{L}_H &= \frac{1}{2}\partial_\mu H \partial^\mu H - \frac{1}{2}m_H^2 H^2 - \upsilon\lambda H^3 - \frac{\lambda}{4}H^4 \\ &= \frac{1}{2}\partial_\mu H \partial^\mu H - \frac{1}{2}m_H^2 H^2 - \frac{m_H^2}{2\upsilon}H^3 - \frac{m_H^2}{8\upsilon^2}H^4 \end{aligned} \tag{8.4.25}$$

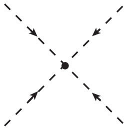
The result $\lambda = m_H^2/(2\upsilon^2)$ has been used in the second step in Eq. (8.4.25), which can be derived from Eqs. (8.4.8) and (8.4.21). The first two terms of Eq. (8.4.25) are the kinetic and mass terms of the Higgs field, respectively. These two terms result in the propagator of Higgs field in the perturbation theory.

$$\cdots\rightarrow\cdots \frac{\mathrm{i}}{k^2 - m_H^2 + \mathrm{i}\varepsilon} \quad \text{(B–D metric)}$$

The last two terms represent the self-interaction of the Higgs field, which give the three-point and four-point vertices in the perturbation theory



$-3\mathrm{i}\dfrac{m_H^2}{\upsilon}$



$-3\mathrm{i}\dfrac{m_H^2}{\upsilon^2}$

8.4.2. *The mass term of gauge field*

The covariant derivative in the Lagrangian in Eq. (8.4.5) that describes the interaction between the scalar and gauge field is

$$\begin{aligned} D_\mu &= \partial_\mu - \mathrm{i}gT^i A_\mu^i - \mathrm{i}\frac{g'}{2}YB_\mu \\ &= \partial_\mu - \frac{\mathrm{i}}{2}\begin{pmatrix} gA_\mu^3 + g'B_\mu & \sqrt{2}gW_\mu^+ \\ \sqrt{2}gW_\mu^- & -gA_\mu^3 + g'B_\mu \end{pmatrix} \end{aligned} \tag{8.4.26}$$

Equation (8.4.26) can be also seen in Eq. (8.3.17), here the weak hypercharge Y=1 for the scalar doublet has been taken. Only substituting the vacuum expectation value of $\Phi(x)$ in Eq. (8.4.24) into Eq. (8.4.5), we can then obtain quadratic term for the gauge field,

$$\begin{aligned}
\mathcal{L}_{\rm GM} &= \left|\left[\partial_\mu - \frac{\rm i}{2}\begin{pmatrix} gA_\mu^3 + g'B_\mu & \sqrt{2}gW_\mu^+ \\ \sqrt{2}gW_\mu^- & -gA_\mu^3 + g'B_\mu \end{pmatrix}\right]\frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ \upsilon \end{pmatrix}\right|^2 \\
&= \left|-\frac{\rm i}{2\sqrt{2}}\begin{pmatrix} gA_\mu^3 + g'B_\mu & \sqrt{2}gW_\mu^+ \\ \sqrt{2}gW_\mu^- & -gA_\mu^3 + g'B_\mu \end{pmatrix}\begin{pmatrix} 0 \\ \upsilon \end{pmatrix}\right|^2 \\
&= \left|-\frac{\rm i}{2\sqrt{2}}\upsilon\begin{pmatrix} \sqrt{2}gW_\mu^+ \\ -gA_\mu^3 + g'B_\mu \end{pmatrix}\right|^2 \\
&= \frac{1}{4}g^2\upsilon^2 W_\mu^+ W^{-\mu} + \frac{1}{8}\upsilon^2\left(g^2 A_\mu^3 A^{3\mu} - 2gg'A_\mu^3 B^\mu + g'^2 B_\mu B^\mu\right) \\
&= \frac{1}{4}g^2\upsilon^2 W_\mu^+ W^{-\mu} + \frac{1}{8}\upsilon^2\left(A_\mu^3, B_\mu\right)\begin{pmatrix} g^2 & -gg' \\ -gg' & g'^2 \end{pmatrix}\begin{pmatrix} A_\mu^3 \\ B^\mu \end{pmatrix}
\end{aligned} \tag{8.4.27}$$

The first term in Eq. (8.4.27) is the mass term of $W^\pm$. We shall see later that the kinetic term of $W^\pm$ is $\partial_\mu W_\nu^+ \partial^\nu W^{-\mu} - \partial_\mu W_\nu^+ \partial^\mu W^{-\nu}$, and the relevant mass term should be in the form $m_{\rm W}^2 W_\mu^+ W^{-\mu}$, therefore

$$m_{\rm W}^2 = \frac{1}{4}g^2\upsilon^2$$

i.e.

$$m_{\rm W} = \frac{1}{2}g\upsilon \tag{8.4.28}$$

We have obtained the relation between the Fermi constant $G_{\rm F}$, $m_{\rm W}$ and g,

$$\frac{G_{\rm F}}{\sqrt{2}} = \frac{g^2}{8m_{\rm W}^2}$$

Substituting Eq. (8.4.28) into the above equation, we can solve the vacuum expectation value of the scalar field

$$\upsilon = \sqrt{\frac{1}{\sqrt{2}G_{\rm F}}} = 246.3\,\text{GeV} \tag{8.4.29}$$

Equation (8.4.29) shows that the vacuum expectation value of the scalar field is 246.3 GeV, which is referred to as the electroweak symmetry breaking scale.

The second term of Eq. (8.4.27) is the mass term for the neutral vector field A_μ^3 and B_μ, the mass matrix of which is non-diagonalized. This shows that A_μ^3 and B_μ are all not the mass eigenstates. The mass eigenvalues of the mass matrix

$$\begin{pmatrix} g^2 & -gg' \\ -gg' & g'^2 \end{pmatrix}$$

can be obtained to be $m_1^2 = g^2 + g'^2$, $m_2^2 = 0$, and the relevant normalized eigenvectors are

$$x_1 = \frac{1}{\sqrt{g^2+g'^2}}\begin{pmatrix} g \\ -g' \end{pmatrix}, \quad x_2 = \frac{1}{\sqrt{g^2+g'^2}}\begin{pmatrix} g' \\ g \end{pmatrix} \tag{8.4.30}$$

According to Eq. (8.4.30), the eigenstates of the neutral vector fields should be

$$\begin{aligned} Z_\mu &= \frac{1}{\sqrt{g^2+g'^2}}(gA_\mu^3 - g'B_\mu) \\ A_\mu &= \frac{1}{\sqrt{g^2+g'^2}}(g'A_\mu^3 + gB_\mu) \end{aligned} \tag{8.4.31}$$

Define

$$\sin\theta = \frac{g'}{\sqrt{g^2+g'^2}}, \quad \cos\theta = \frac{g}{\sqrt{g^2+g'^2}} \tag{8.4.32}$$

where θ is referred to as the Weinberg angle. With Weinberg angle, Eq. (8.4.31) can be expressed as

$$\begin{aligned} A_\mu &= \cos\theta B_\mu + \sin\theta A_\mu^3 \\ Z_\mu &= -\sin\theta B_\mu + \cos\theta A_\mu^3 \end{aligned} \tag{8.4.33}$$

which has been used in Sec. 8.3 as Eq. (8.3.18). We can solve A_μ^3 and B_μ from Eq. (8.4.33),

$$\begin{aligned} A_\mu^3 &= \sin\theta A_\mu + \cos\theta Z_\mu \\ B_\mu &= \cos\theta A_\mu - \sin\theta Z_\mu \end{aligned} \tag{8.4.34}$$

Substituting Eq. (8.4.34) into Eq. (8.4.27), we can obtained the mass term for gauge field

$$\mathcal{L}_{\mathrm{GM}} = \frac{1}{4}g^2\upsilon^2 W_\mu^+ W^{-\mu} + \frac{1}{8}\upsilon^2(g^2 + g'^2)Z_\mu Z^\mu \tag{8.4.35}$$

Equation (8.4.35) indicates that particle Z has obtained mass, while photon is still massless. Later we will derive that the kinetic energy term of Z_μ should be $\frac{1}{2}(\partial_\mu Z_\nu \partial^\nu Z^\mu - \partial_\mu Z_\nu \partial^\mu Z^\nu)$, and the mass term relevant to such kinetic energy is $\frac{1}{2}m_Z^2 Z_\mu Z^\mu$. Therefore from Eq. (8.4.35), we can obtain

$$\frac{1}{8}\upsilon^2(g^2 + g'^2) = \frac{1}{2}m_{\mathrm{Z}}^2$$

That is

$$m_{\mathrm{Z}} = \frac{1}{2}\upsilon\sqrt{g^2 + g'^2} \tag{8.4.36}$$

Comparing Eqs. (8.4.28) and (8.4.36), we can obtain the relation between particles w and z,

$$\frac{m_{\mathrm{W}}}{m_{\mathrm{Z}}} = \cos\theta \tag{8.4.37}$$

In addition, considering Eqs. (8.4.29) and (8.3.24), we can get

$$\begin{aligned} m_{\mathrm{W}} &= \frac{1}{2}g\upsilon = \frac{1}{2}\frac{e}{\sin\theta}\upsilon = \frac{1}{2}\sqrt{\frac{4\pi\alpha}{\sqrt{2}\sin^2\theta G_{\mathrm{F}}}} \\ &\approx 78\,\mathrm{GeV} \end{aligned} \tag{8.4.38}$$

and

$$m_{\mathrm{Z}} = \frac{m_{\mathrm{W}}}{\cos\theta} \approx 89\,\mathrm{GeV} \tag{8.4.39}$$

The above two equations are the predictions about the masses of the intermediate bosons W and Z in Glashow–Weinberg–Salam electroweak unification model. In the above results, the effect of higher-order corrections in perturbation theory and the renormalization effect are not considered. After considering these effects, the theoretical prediction will slightly increase, which in good agreement with the results of the later experimental measurement.

8.4.3. *The interaction term of the scalar and gauge fields*

In $\mathcal{L}_H$ of Eq. (8.4.5), except the part for the pure scalar field and the pure gauge field, the other items describe the interaction between the scalar and gauge field. Substituting Eq. (8.4.34) into Eq. (8.4.26) and considering the relation between the Weinberg angle and couplings g and g', we can express the covariant derivative in terms of the mass eigenstates of the gauge fields,

$$D_\mu = \partial_\mu - \mathrm{i}\begin{pmatrix} eA_\mu + g\dfrac{\cos 2\theta}{2\cos\theta}Z_\mu & \dfrac{g}{\sqrt{2}}W_\mu^+ \\ \dfrac{g}{\sqrt{2}}W_\mu^- & -\dfrac{g}{2\cos\theta}Z_\mu \end{pmatrix} \tag{8.4.40}$$

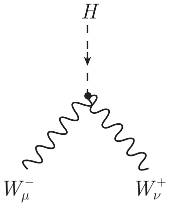
then

$$\begin{aligned}(D_\mu\Phi)^+ D^\mu\Phi &= \left|\left[\partial_\mu - \mathrm{i}\begin{pmatrix} eA_\mu + g\dfrac{\cos 2\theta}{2\cos\theta}Z_\mu & \dfrac{g}{\sqrt{2}}W_\mu^+ \\ \dfrac{g}{\sqrt{2}}W_\mu^- & -\dfrac{g}{2\cos\theta}Z_\mu \end{pmatrix}\right]\right. \\ &\qquad \left.\times\frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ \upsilon + H\end{pmatrix}\right|^2 \\ &= \frac{1}{2}\left|\begin{pmatrix} -\mathrm{i}\dfrac{g}{\sqrt{2}}W_\mu^+(\upsilon+H) \\ \partial_\mu H + \mathrm{i}\dfrac{g}{2\cos\theta}Z_\mu(\upsilon+H)\end{pmatrix}\right|^2 \\ &= \frac{1}{4}g^2 W_\mu^+ W^{-\mu}(\upsilon+H)^2 \\ &\quad + \frac{1}{2}\partial_\mu H\partial^\mu H + \frac{g^2}{8\cos^2\theta}Z_\mu Z^\mu(\upsilon+H)^2 \end{aligned} \tag{8.4.41}$$

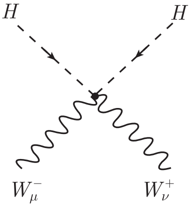
where the interaction terms between the scalar and gauge fields are

$$\begin{aligned}\mathcal{L}_{\mathrm{HG}} &= \frac{1}{2}g^2\upsilon W_\mu^+ W^{-\mu}H + \frac{1}{4}g^2 W_\mu^+ W^{-\mu}H^2 + \frac{g^2\upsilon}{4\cos^2\theta}Z_\mu Z^\mu H \\ &\quad + \frac{g^2}{8\cos^2\theta}Z_\mu Z^\mu H^2 \end{aligned} \tag{8.4.42}$$

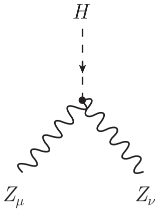
According to the perturbation theory, the Feynman rules relevant to these terms are



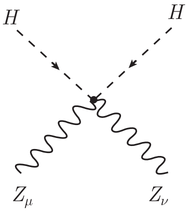
$$\mathrm{i}\frac{1}{2}g^2\upsilon g_{\mu\nu}=\mathrm{i}gm_w\,g_{\mu\nu}$$



$$\mathrm{i}\frac{1}{2}g^2\,g_{\mu\nu}$$



$$\mathrm{i}\frac{g^2\upsilon}{2\cos^2\theta}g_{\mu\nu}=\mathrm{i}\frac{2m_Z^2}{\upsilon}\,g_{\mu\nu}$$



$$\mathrm{i}\frac{g^2}{2\cos^2\theta}\,g_{\mu\nu}=\mathrm{i}\frac{2m_Z^2}{\upsilon^2}g_{\mu\nu}$$

8.5. The Gain of the Fermion Mass: Yukawa Coupling

Fermions can also interact with the scalar doublet $\Phi(x)$ in a way that keeps local gauge symmetry. The coupling can be written similar to the coupling of nucleons and pion, which is called Yukawa coupling.

For the lepton field, we need to consider the left-handed doublet and right-handed singlet of the lepton field

$$\begin{aligned}&l_{\mathrm{L}}^i=\begin{pmatrix}\nu_{\mathrm{eL}}\\ \mathrm{e}_{\mathrm{L}}\end{pmatrix},\quad\begin{pmatrix}\nu_{\mu\mathrm{L}}\\ \mu_{\mathrm{L}}\end{pmatrix},\quad\begin{pmatrix}\nu_{\tau\mathrm{L}}\\ \tau_{\mathrm{L}}\end{pmatrix}\\ &l_{i\mathrm{R}}=\mathrm{e}_{\mathrm{R}},\quad\mu_{\mathrm{R}},\quad\tau_{\mathrm{R}}\end{aligned}\tag{8.5.1}$$

here it is assumed that there is no right-handed neutrino in nature. For the quark field, it is necessary to consider the quark left-handed doublet and right-handed singlet

$$\begin{aligned}&q_{\mathrm{L}}^i=\begin{pmatrix}u_{\mathrm{L}}\\ d'_{\mathrm{L}}\end{pmatrix},\quad\begin{pmatrix}c_{\mathrm{L}}\\ s'_{\mathrm{L}}\end{pmatrix},\quad\begin{pmatrix}t_{\mathrm{L}}\\ b'_{\mathrm{L}}\end{pmatrix}\\ &u_{i\mathrm{R}}=u_{\mathrm{R}},\quad c_{\mathrm{R}},\quad t_{\mathrm{R}}\\ &d'_{i\mathrm{R}}=d'_{\mathrm{R}},\quad s'_{\mathrm{R}},\quad b'_{\mathrm{R}}\end{aligned}\tag{8.5.2}$$

where d_i' denotes weak interaction eigenstate, while the mass eigenstate is denoted by d_i. Equations (8.5.1) and (8.5.2) indicate that only the down-type quarks mix, while up-type quarks and leptons do not mix

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = U \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \quad U^+U = 1 \tag{8.5.3}$$

However, the more general case is that both up-type quarks, down-type quarks and leptons can mix, but in practice this is not necessary. For quarks, no matter mixing occurs between up-type quarks or down-type quarks, both eventually lead to the same physical effect for weak interaction, therefore only one case needs to be considered. Let us illustrate this issue with a charged current.

We start from the case that both up- and down-type quarks can mix. Let

$$u_i' = P_{ij} u_j, \quad d_i' = R_{ij} d_j \tag{8.5.4}$$

where P and R are the unitary transformation matrices, which satisfy

$$P^+P = 1, \quad R^+R = 1$$

Then the charge current

$$\begin{aligned} \bar{u}_i' \hat{O} d_i' &= \bar{u}_j (P^+)_{ji} \hat{O} R_{ik} d_k \\ &= (P^+)_{ji} R_{ik} \bar{u}_j \hat{O} d_k \\ &= V_{jk} \bar{u}_j \hat{O} d_k \end{aligned} \tag{8.5.5}$$

where $\hat{O}$ represents any operator in spinor space, and $V_{jk} = (P^+)_{ji} R_{ik}$. On the other hand, for the case that mixing only occurs in down-type quarks, let

$$d_i' = U_{ij} d_j \tag{8.5.6}$$

and the weak charged current is

$$\begin{aligned} \bar{u}_i \hat{O} d_i' &= \bar{u}_i \hat{O} U_{ik} d_k \\ &= U_{ik} \bar{u}_i \hat{O} d_k \\ &= U_{jk} \bar{u}_j \hat{O} d_k \end{aligned} \tag{8.5.7}$$

For the results in Eqs. (8.5.5) and (8.5.7), the quantity that can be measured in experiment is V_{jk} or U_{jk}, while P_{ij} and R_{ij} are not direct observable. For the measurement in the same physical process, V and U should be the same quantities. Therefore Eqs. (8.5.5) and (8.5.7) are equivalent in physics. Similarly, we can also verify that the cases of only up-type quark or down-type quark mixing lead to the same physical result.

For leptons, let the mixing of neutrinos and charged leptons be

$$\nu_i' = P_{ij}'\nu_j, \quad e_i' = R_{ij}'e_i, \tag{8.5.8}$$

respectively, then the weak charged current for leptons is

$$\begin{aligned}\bar{\nu}_i'\hat{O}e_i' &= \bar{\nu}_j(P'^{+})_{ji}\hat{O}R_{ik}'e_k \\ &= (P'^{+})_{ji}R_{ik}'\bar{\nu}_j\hat{O}e_k\end{aligned} \tag{8.5.9}$$

If only consider the case of massless neutrinos, then these three types of neutrinos can not be distinguished by masses, and they can only be distinguished by the associated leptons that interact with them in weak interaction. Therefore, the mixing of massless neutrinos can be arbitrary, and we can take

$$P' = R' \tag{8.5.10}$$

According to the unitarity of P' and R', we have

$$(P'^{+})_{ji}R_{ik}' = \delta_{jk} \tag{8.5.11}$$

Then Eq. (8.5.9) becomes

$$\bar{\nu}_i'\hat{O}e_i' = \bar{\nu}_j\hat{O}e_j \tag{8.5.12}$$

That is, if neutrinos are massless, then it is equivalent no matter leptons can mix or not.

Next we shall discuss how to construct the gauge invariant interaction between fields of lepton, quark and scalar doublet in Eqs. (8.5.1) and (8.5.2). The following combination

$$(\Phi^{+}l_{\mathrm{L}}^{i})$$

is SU(2) gauge invariant. But such a term includes the fermion field l_{L}^{i} only one time, so it is not a Dirac scalar in spinor space, i.e. it includes Dirac matrix, while Lagrangian must be a complex number. To construct a C

number with the combination $(\Phi^+ l_\mathrm{L}^i)$ which is SU(2) gauge invariant, we can multiply it with $\bar{l}_{i\mathrm{R}}$ from the left

$$\bar{l}_{i\mathrm{R}}\Phi^+ l_\mathrm{L}^i$$

The above combination is still SU(2) gauge invariant, because $\bar{l}_{i\mathrm{R}}$ is a singlet under SU(2) transformation.

Next let us check the $\mathrm{U(1)_Y}$ transformation of $(\bar{l}_{i\mathrm{R}}\Phi^+ l_\mathrm{L}^i)$. The weak hypercharge of the left-handed lepton is $Y_1 = -1$, and for right-handed lepton $Y_2 = -2$. Denote the hypercharge of the scalar doublet as Y, then under the $\mathrm{U(1)_Y}$ transformation,

$$\begin{aligned} l_\mathrm{L}^{\prime i} &= \mathrm{e}^{\mathrm{i}\dfrac{g'}{2}Y_1\alpha} l_\mathrm{L}^i \\ l'_{i\mathrm{R}} &= \mathrm{e}^{\mathrm{i}\dfrac{g'}{2}Y_2\alpha} l_{i\mathrm{R}} \\ \Phi' &= \mathrm{e}^{\mathrm{i}\frac{g'}{2}Y\alpha}\Phi \end{aligned}$$

the transformation of $(\bar{l}_{i\mathrm{R}}\Phi^+ l_\mathrm{L}^i)$ is

$$\begin{aligned} \bar{l}'_{i\mathrm{R}}\Phi'^+ l_\mathrm{L}^{\prime i} &= \bar{l}_{\mathrm{iR}}\mathrm{e}^{-\mathrm{i}\frac{g'}{2}Y_2\alpha}\Phi^+\mathrm{e}^{-\mathrm{i}\frac{g'}{2}Y\alpha}\mathrm{e}^{\mathrm{i}\frac{g'}{2}Y_1\alpha} l_\mathrm{L}^i \\ &= \mathrm{e}^{\mathrm{i}\frac{g'}{2}(-Y_2 - Y + Y_1)\alpha}\bar{l}_{i\mathrm{R}}\Phi^+ l_\mathrm{L}^i \end{aligned} \qquad (8.5.13)$$

One can find that, if the hypercharge of the scalar doublet is

$$Y = Y_1 - Y_2 = (-1) - (-2) = 1 \qquad (8.5.14)$$

then $(\bar{l}_{i\mathrm{R}}\Phi^+ l_\mathrm{L}^i)$ is also $\mathrm{U(1)_Y}$ invariant, therefore it is also $\mathrm{SU(2)_L \times U(1)_Y}$ invariant. If the hypercharge is not taken as Eq. (8.5.14), the interaction of the scalar and fermion will be in a more complicated form. So for the principle of simplicity, we take the weak hypercharge of the scalar field double to be Y = 1.

Similarly, the gauge invariant interaction of quark and scalar fields can also be constructed. Therefore, including the Yukawa coupling of leptons and scalar field, the total Yukawa couplings are

$$\mathcal{L}_\mathrm{Yukawa} = -g_i^\mathrm{L}\bar{l}_{i\mathrm{R}}\Phi^+ l_\mathrm{L}^i - g_{ij}^d \bar{d}'_{i\mathrm{R}}\Phi^+ q_\mathrm{L}^j - g_i^u \bar{u}_{i\mathrm{R}}\tilde{\Phi}^+ q_\mathrm{L}^i + \mathrm{h.c.} \qquad (8.5.15)$$

where the first term is coupling for leptons, the second and third terms are couplings involving quarks, and h.c. stands for Hermitian terms, g_i^L, g_{ij}^d,

g_i^u are Yukawa couplings. The way of taking i and $j = 1, 2, 3$ in he first and third terms means not considering the mixing of leptons and up-type quarks, and only the mixing of down-type quarks is considered. In Eq. (8.5.15),

$$\begin{aligned}\tilde{\Phi} = \mathrm{i}\sigma_2\Phi^* &= \mathrm{i}\begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}\begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}^* \\ &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} \phi^- \\ \phi^{0*} \end{pmatrix} = \begin{pmatrix} \phi^{0*} \\ -\phi^- \end{pmatrix} \end{aligned} \tag{8.5.16}$$

The reason why the scalar field coupling with the right-handed up-type quarks is taken to be $\tilde{\Phi}$ is for up-type quarks gaining mass after spontaneous symmetry breaking. In addition, using Eq. (4.1.30), we can verify that the SU(2) transformation of $\tilde{\Phi}$ is the same as that of Φ, i.e.

$$\tilde{\Phi} \to \tilde{\Phi}' = U\tilde{\Phi} = \mathrm{e}^{\mathrm{i}gT^j\beta^j}\tilde{\Phi} \tag{8.5.17}$$

Thus, the third term of Eq. (8.5.15) is guaranteed to be gauge invariant. The first and second terms and their Hermitian in Eq. (8.5.15) will lead to leptons and down-type quarks obtaining masses after symmetry breaking, respectively.

Substituting the expressions of $\Phi(x)$ and $\tilde{\Phi}$ after symmetry breaking

$$\Phi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ \upsilon + H(x) \end{pmatrix}, \quad \tilde{\Phi}(x) = \frac{1}{\sqrt{2}}\begin{pmatrix} \upsilon + H(x) \\ 0 \end{pmatrix} \tag{8.5.18}$$

into $\mathcal{L}_{\text{Yukawa}}$ in Eq. (8.5.15), we can obtain

$$\begin{aligned}\mathcal{L}_{\text{Yukawa}} &= -\frac{g_i^{\mathrm{L}}}{\sqrt{2}}\bar{l}_{i\mathrm{R}}(0 \quad \upsilon + H)\begin{pmatrix} \nu_{i\mathrm{L}} \\ l_{i\mathrm{L}} \end{pmatrix} - \frac{g_{ij}^d}{\sqrt{2}}\bar{d}'_{i\mathrm{R}}(0 \quad \upsilon + H)\begin{pmatrix} u_{j\mathrm{L}} \\ d'_{j\mathrm{L}} \end{pmatrix} \\ &\quad - \frac{g_i^u}{\sqrt{2}}\bar{u}_{i\mathrm{R}}(\upsilon + H \quad 0)\begin{pmatrix} u_{i\mathrm{L}} \\ d'_{i\mathrm{L}} \end{pmatrix} + \text{h.c.} \\ &= -\frac{g_i^{\mathrm{L}}}{\sqrt{2}}\bar{l}_{i\mathrm{R}}l_{i\mathrm{L}}(\upsilon + H) - \frac{g_{ij}^d}{\sqrt{2}}\bar{d}'_{i\mathrm{R}}d'_{j\mathrm{L}}(\upsilon + H) \\ &\quad - \frac{g_i^u}{\sqrt{2}}\bar{u}_{i\mathrm{R}}u_{i\mathrm{L}}(\upsilon + H) + \text{h.c.}\end{aligned} \tag{8.5.19}$$

The matrix (g^d_{ij}) in Eq. (8.5.19) is non-diagonal, which implies that d' is not its mass eigenstate. Mathematically, for any $n \times n$ matrix M, there are always $n \times n$ unitary matrices U and V, so that U^+MV is diagonalized, and the phase angle of U and V can be further appropriately selected, so that the elements of the diagonalized matrix are all not less than 0.

Let P^d_{R} and P^d_{L} be the unitary matrices that can diagonalize the matrix (g^d_{ij}), i.e.

$$(P^{d+}_{\mathrm{R}})_{ki} g^d_{ij} (P^d_{\mathrm{L}})_{jl} = g^d_k \delta_{kl} \tag{8.5.20}$$

Note that there is no summation for the repeating indices at the right-hand side of the equation, and g^d_k, $k = 1, 2, 3$, stands for the diagonal elements after (g^d_{ij}) being diagonalized. Perform transformation for the down-type quarks

$$\begin{aligned} d'_{i\mathrm{R}} &= (P^d_{\mathrm{R}})_{ik} d_{k\mathrm{R}} \\ d'_{j\mathrm{L}} &= (P^d_{\mathrm{L}})_{jl} d_{l\mathrm{L}} \end{aligned} \tag{8.5.21}$$

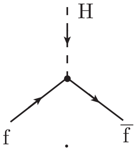
then Eq. (8.5.19) becomes

$$\begin{aligned} \mathcal{L}_{\text{Yukawa}} &= -\frac{g^{\mathrm{L}}_i}{\sqrt{2}} \bar{l}_{i\mathrm{R}} l_{i\mathrm{L}} (\upsilon + H) - \frac{g^d_i}{\sqrt{2}} \bar{d}_{i\mathrm{R}} d_{i\mathrm{L}} (\upsilon + H) \\ &\quad - \frac{g^u_i}{\sqrt{2}} \bar{u}_{i\mathrm{R}} u_{i\mathrm{L}} (\upsilon + H) + \text{h.c.} \\ &= -\frac{g^{\mathrm{L}}_i}{\sqrt{2}} \bar{l}_i l_i (\upsilon + H) - \frac{g^d_i}{\sqrt{2}} \bar{d}_i d_i (\upsilon + H) - \frac{g^u_i}{\sqrt{2}} \bar{u}_i u_i (\upsilon + H) \\ &= -\left[\frac{m_{li}}{\upsilon} \, \bar{l}_i l_i + \frac{m_{di}}{\upsilon} \bar{d}_i d_i + \frac{m_{ui}}{\upsilon} \bar{u}_i u_i \right] H \\ &\quad - m_{li} \bar{l}_i l_i - m_{di} \bar{d}_i d_i - m_{ui} \bar{u}_i u_i \end{aligned} \tag{8.5.22}$$

where the definitions of the left and right-handed fields have been used in the second step, and the Hermitian conjugation has been considered, and

$$m_{li} \equiv \frac{g^{\mathrm{L}}_i}{\sqrt{2}} \upsilon, \quad m_{di} \equiv \frac{g^d_i}{\sqrt{2}} \upsilon, \quad m_{ui} \equiv \frac{g^u_i}{\sqrt{2}} \upsilon \tag{8.5.23}$$

The first three terms in Eq. (8.5.22) are the couplings of leptons and quarks with Higgs field. The final three terms are the mass terms for fermions, which are obtained by spontaneous symmetry breaking. The first three terms lead to the Feynman rule for the coupling between fermion and Higgs fields



$$-\frac{\mathrm{i}}{v}m_{\mathrm{f}}$$

where f denotes any types of fermions.

8.6. The Interaction of Fermion and Gauge Field II

Using the unitary transformation of Eq. (8.5.21), the down-type quarks are transformed from weak eigenstate to their mass eigenstate $d' \to d$. The down-type quarks in Eq. (8.3.25) are all transformed to their mass eigenstate, then the kinetic term is transformed to be

$$\begin{aligned}\bar{d}'_i \mathrm{i}\gamma^\mu \partial_\mu d'_i &= \bar{d}'_{i\mathrm{R}} \mathrm{i}\gamma^\mu \partial_\mu d'_{i\mathrm{R}} + \bar{d}'_{i\mathrm{L}} \mathrm{i}\gamma^\mu \partial_\mu d'_{i\mathrm{L}} \\ &= \bar{d}_{i\mathrm{R}} \mathrm{i}\gamma^\mu \partial_\mu d_{i\mathrm{R}} + \bar{d}_{i\mathrm{L}} \mathrm{i}\gamma^\mu \partial_\mu d_{i\mathrm{L}} \\ &= \bar{d}_i \mathrm{i}\gamma^\mu \partial_\mu d_i \qquad (8.6.1)\end{aligned}$$

where unitarity of P^d_{R} and P^d_{L} has been used in the second step. The other neutral current terms are also transformed by doing the substitution of $d'_i \to d_i$ according to Eq. (8.5.21), and the charged current changes in the following form

$$\begin{aligned}\bar{d}'_i \gamma^\mu (1-\gamma_5) u_i &= \bar{d}'_{i\mathrm{L}} \gamma^\mu (1-\gamma_5) u_i \\ &= (P^{d+}_{\mathrm{L}})_{ji} \bar{d}_{j\mathrm{L}} \gamma^\mu (1-\gamma_5) u_i \\ &= (P^{d+}_{\mathrm{L}})_{ji} \bar{d}_j \gamma^\mu (1-\gamma_5) u_i \qquad (8.6.2)\end{aligned}$$

Define a matrix

$$V \equiv P^d_{\mathrm{L}} \qquad (8.6.3)$$

then Eq. (8.6.2) becomes

$$\bar{d}'_i \gamma^\mu (1-\gamma_5) u_i = V^*_{ij} \bar{d}_j \gamma^\mu (1-\gamma_5) u_i \qquad (8.6.4)$$

Here V_{ij}^* is the complex conjugate of the (i,j)th element of the matrix V. Taking the Hermitian conjugation for both two sides of Eq. (8.6.4), then we can get

$$\bar{u}_i\gamma^\mu(1-\gamma_5)d_i' = V_{ij}\bar{u}_i\gamma^\mu(1-\gamma_5)d_j \tag{8.6.5}$$

The matrix V is referred to as Cabibbo–Kobayashi–Maskawa (CKM) matrix.

Using the transformations in Eqs. (8.6.1)–(8.6.5), the fermion and gauge field interaction terms in Eq. (8.3.25) when expressed in terms of the mass eigenstate of each field become

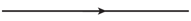
$$\begin{aligned}
\mathcal{L}_{\mathrm{FG}} = {} & \bar{\nu}_{i\mathrm{L}}\mathrm{i}\gamma^\mu\partial_\mu\nu_{i\mathrm{L}} + \bar{l}_i\mathrm{i}\gamma^\mu\partial_\mu l_i + \bar{u}_i\mathrm{i}\gamma^\mu\partial_\mu u_i + \bar{d}_i\mathrm{i}\gamma^\mu\partial_\mu d_i \\
& + \frac{g}{2\sqrt{2}}\left[\bar{l}_i\gamma^\mu\left(1-\gamma_5\right)\nu_i + V_{ij}^*\bar{d}_j\gamma^\mu\left(1-\gamma_5\right)u_i\right]W_\mu^- \\
& + \frac{g}{2\sqrt{2}}\left[\bar{\nu}_i\gamma^\mu\left(1-\gamma_5\right)l_i + V_{ij}\bar{u}_i\gamma^\mu\left(1-\gamma_5\right)d_j\right]W_\mu^+ \\
& + \frac{g}{4\cos\theta}\Bigg[\bar{\nu}_i\gamma^\mu\left(1-\gamma_5\right)\nu_i + \bar{l}_i\gamma^\mu\left(-(1-4\sin^2\theta)+\gamma_5\right)l_i \\
& + \bar{u}_i\gamma^\mu\left(\left(1-\frac{8}{3}\sin^2\theta\right)-\gamma_5\right)u_i + \bar{d}_i\gamma^\mu \\
& \times\left(-\left(1-\frac{4}{3}\sin^2\theta\right)+\gamma_5\right)d_i\Bigg]Z_\mu \\
& + e\left[-\bar{l}_i\gamma^\mu l_i + \frac{2}{3}\bar{u}_i\gamma^\mu u_i - \frac{1}{3}\bar{d}_i\gamma^\mu d_i\right]A_\mu
\end{aligned} \tag{8.6.6}$$

Add Yukawa coupling terms to Eq. (8.6.6), then the total Lagrangian about fermions is

$$\mathcal{L}_{\mathrm{F}} = \mathcal{L}_{\mathrm{FG}} + L_{\mathrm{Yukawa}} \tag{8.6.7}$$

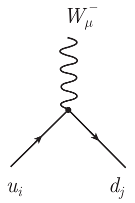
Equation (8.6.7) can give Feynman rules about fermions.

The propagator of fermion:

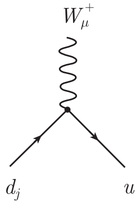


$$\frac{\mathrm{i}}{\not{p}-m+\mathrm{i}\varepsilon} = \frac{\mathrm{i}(\not{p}+m)}{p^2-m^2+\mathrm{i}\varepsilon}$$

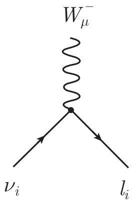
The interaction vertex of fermions and gauge field:



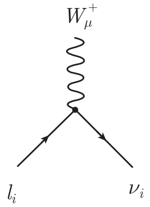
$$\mathrm{i}\,\frac{g}{2\sqrt{2}}\,V_{ij}^*\gamma^\mu(1-\gamma_5)$$



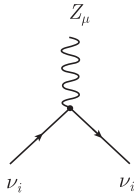
$$\mathrm{i}\,\frac{g}{2\sqrt{2}}\,V_{ij}\gamma^\mu(1-\gamma_5)$$



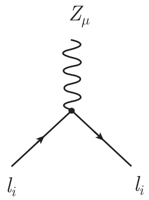
$$\mathrm{i}\,\frac{g}{2\sqrt{2}}\,\gamma^\mu(1-\gamma_5)$$



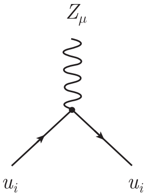
$$\mathrm{i}\,\frac{g}{2\sqrt{2}}\,\gamma^\mu(1-\gamma_5)$$



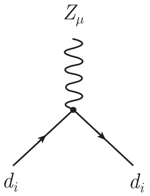
$$\mathrm{i}\,\frac{g}{4\cos\,\theta}\,\gamma^\mu(1-\gamma_5)$$



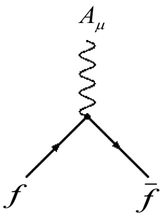
$$\mathrm{i}\,\frac{g}{4\cos\,\theta}\,\gamma^\mu\cdot$$
$$[-(1-4\sin^2\,\theta)+\gamma_5]$$



$$\mathrm{i}\,\frac{g}{4\cos\,\theta}\,\gamma^\mu\cdot$$
$$\left[\left(1-\frac{8}{3}\sin^2\theta\right)-\gamma_5\right]$$



$$\mathrm{i}\,\frac{g}{4\cos\,\theta}\,\gamma^\mu\cdot$$
$$\left[-\left(1-\frac{4}{3}\sin^2\theta\right)+\gamma_5\right]$$



f denotes fermion

$$ieQ_f\gamma^\mu$$

8.7. The Interaction of Gauge Fields

It has been proved previously that gauge invariant interaction can be made up with pure gauge fields, and the Lagrangian can be written in the following form:

$$\mathcal{L}_{\mathrm{G}} = -\frac{1}{4}F^i_{\mu\nu}F^{i\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \tag{8.7.1}$$

where $F^i_{\mu\nu}$, $i = 1, 2, 3$, is the SU(2) gauge field tensor, and $B_{\mu\nu}$ the U(1) gauge field tensor

$$\begin{aligned} F^i_{\mu\nu} &= \partial_\mu A^i_\nu - \partial_\nu A^i_\mu + g\varepsilon^{ijk}A^j_\mu A^k_\nu \\ B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu \end{aligned} \tag{8.7.2}$$

The gauge fields A^i_μ and B_μ all are not the mass eigenstates in the above equations. We shall express the Lagrangian in terms of the mass eigenstates of the gauge fields in the following, in order that the quanta of these gauge fields are directly physical particles. According to previous discussion, the mass eigenstates are $W^\pm_\mu$, Z_μ and the electromagnetic field A_μ. The relations of these mass eigenstates and the fields A^i_μ, B_μ are

$$\begin{aligned} A^1_\mu &= \frac{1}{\sqrt{2}}(W^+_\mu + W^-_\mu) \\ A^2_\mu &= \frac{\mathrm{i}}{\sqrt{2}}(W^+_\mu - W^-_\mu) \end{aligned} \tag{8.7.3}$$

and

$$\begin{aligned} A^3_\mu &= \sin\theta A_\mu + \cos\theta Z_\mu \\ B_\mu &= \cos\theta A_\mu - \sin\theta Z_\mu \end{aligned} \tag{8.7.4}$$

Substituting Eqs. (8.7.3) and (8.7.4) into Eq. (8.7.1), we can then obtain the Lagrangian $\mathcal{L}_{\mathrm{G}}$ that is expressed in terms the mass eigenstates of the gauge fields. Let us derive the expression of $F^i_{\mu\nu}$ and $B_{\mu\nu}$ in terms of the mass eigenstates of the gauge fields at first

$$\begin{aligned} F^1_{\mu\nu} &= \partial_\mu A^1_\nu - \partial_\nu A^1_\mu + g\varepsilon^{1jk}A^j_\mu A^k_\nu \\ &= \partial_\mu A^1_\nu - \partial_\nu A^1_\mu + g(A^2_\mu A^3_\nu - A^3_\mu A^2_\nu) \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{\sqrt{2}}(\partial_\mu W_\nu^+ + \partial_\mu W_\nu^- - \partial_\nu W_\mu^+ - \partial_\nu W_\mu^-) + \frac{\mathrm{i}}{\sqrt{2}} g \left[(W_\mu^+ - W_\mu^-)\right. \\
&\quad \times (\sin\theta A_\nu + \cos\theta Z_\nu) - (\sin\theta A_\mu + \cos\theta Z_\mu)\left.(W_\nu^+ - W_\nu^-)\right] \\
&= \frac{1}{\sqrt{2}}(\partial_\mu W_\nu^+ - \partial_\nu W_\mu^+ + \partial_\mu W_\nu^- - \partial_\nu W_\mu^-) \\
&\quad + \frac{\mathrm{i}}{\sqrt{2}} g \left[\sin\theta[(W_\mu^+ - W_\mu^-)\, A_\nu - A_\mu(W_\nu^+ - W_\nu^-)]\right. \\
&\quad \left. + \cos\theta[(W_\mu^+ - W_\mu^-)Z_\nu - Z_\mu(W_\nu^+ - W_\nu^-)]\right]
\end{aligned}
\tag{8.7.5a}
$$

$$
\begin{aligned}
F_{\mu\nu}^2 &= \partial_\mu A_\nu^2 - \partial_\nu A_\mu^2 + g\varepsilon^{2jk} A_\mu^j A_\nu^k \\
&= \partial_\mu A_\nu^2 - \partial_\nu A_\mu^2 + g(A_\mu^3 A_\nu^1 - A_\mu^1 A_\nu^3) \\
&= \frac{\mathrm{i}}{\sqrt{2}}(\partial_\mu W_\nu^+ - \partial_\mu W_\nu^- - \partial_\nu W_\mu^+ + \partial_\nu W_\mu^-) + \frac{g}{\sqrt{2}}\left[(\sin\theta A_\mu + \cos\theta Z_\mu)\right. \\
&\quad \times (W_\nu^+ + W_\nu^-) - (W_\mu^+ + W_\mu^-)\left.(\sin\theta A_\nu + \cos\theta Z_\nu)\right] \\
&= \frac{\mathrm{i}}{\sqrt{2}}\left[\partial_\mu W_\nu^+ - \partial_\nu W_\mu^+ - (\partial_\mu W_\nu^- - \partial_\nu W_\mu^-)\right] \\
&\quad + \frac{g}{\sqrt{2}}\left[\sin\theta[A_\mu(W_\nu^+ + W_\nu^-) - (W_\mu^+ + W_\mu^-)A_\nu]\right. \\
&\quad \left. + \cos\theta[Z_\mu(W_\nu^+ + W_\nu^-) - (W_\mu^+ + W_\mu^-)Z_\nu]\right]
\end{aligned}
\tag{8.7.5b}
$$

$$
\begin{aligned}
F_{\mu\nu}^3 &= \partial_\mu A_\nu^3 - \partial_\nu A_\mu^3 + g\varepsilon^{3jk} A_\mu^j A_\nu^k \\
&= \partial_\mu A_\nu^3 - \partial_\nu A_\mu^3 + g(A_\mu^1 A_\nu^2 - A_\mu^2 A_\nu^1) \\
&= \partial_\mu(\sin\theta A_\nu + \cos\theta Z_\nu) - \partial_\nu(\sin\theta A_\mu + \cos\theta Z_\mu) \\
&\quad + \frac{\mathrm{i}}{2} g\left[(W_\mu^+ + W_\mu^-)\,(W_\nu^+ - W_\nu^-) - (W_\mu^+ - W_\mu^-)(W_\nu^+ + W_\nu^-)\right] \\
&= \sin\theta(\partial_\mu A_\nu - \partial_\nu A_\mu) + \cos\theta(\partial_\mu Z_\nu - \partial_\nu Z_\mu) \\
&\quad - \mathrm{i}g(W_\mu^+ W_\nu^- - W_\mu^- W_\nu^+)
\end{aligned}
\tag{8.7.5c}
$$

$$
\begin{aligned}
B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu \\
&= \partial_\mu(\cos\theta A_\nu - \sin\theta Z_\nu) - \partial_\nu(\cos\theta A_\mu - \sin\theta Z_\mu) \\
&= \cos\theta(\partial_\mu A_\nu - \partial_\nu A_\mu) - \sin\theta(\partial_\mu Z_\nu - \partial_\nu Z_\mu)
\end{aligned}
\tag{8.7.5d}
$$

For simplicity in the calculation, we define two auxiliary quantities at first instead of substituting Eqs. (8.7.5a)–(8.7.5d) into Eq. (8.7.1) directly. The auxiliary quantities are

$$F^+_{\mu\nu} = \frac{1}{\sqrt{2}}(F^1_{\mu\nu} - \mathrm{i}F^2_{\mu\nu}), \quad F^-_{\mu\nu} = \frac{1}{\sqrt{2}}(F^1_{\mu\nu} + \mathrm{i}F^2_{\mu\nu}) \tag{8.7.6}$$

Using Eqs. (8.7.5a) and (8.7.5b), we can obtain

$$\begin{aligned} F^+_{\mu\nu} = \partial_\mu W^+_\nu - \partial_\nu W^+_\mu &+ \mathrm{i}g[\sin\theta(W^+_\mu A_\nu - W^+_\nu A_\mu) \\ &+ \cos\theta(W^+_\mu Z_\nu - W^+_\nu Z_\mu)] \end{aligned} \tag{8.7.7a}$$

$$\begin{aligned} F^-_{\mu\nu} = \partial_\mu W^-_\nu - \partial_\nu W^-_\mu &- \mathrm{i}g\left[\sin\theta(W^-_\mu A_\nu - W^-_\nu A_\mu)\right. \\ &\left. + \cos\theta(W^-_\mu Z_\nu - W^-_\nu Z_\mu)\right] \end{aligned} \tag{8.7.7b}$$

The above two equations are Hermitian conjugate to each other.

Using the definition of $F^\pm_{\mu\nu}$, we can prove

$$\sum_{i=1}^{2} F^i_{\mu\nu}F^{i\mu\nu} = 2F^+_{\mu\nu}F^{-\mu\nu} \tag{8.7.8}$$

Then using Eq. (8.7.8) and (8.7.7), and with a few steps of calculation, we can obtain

$$\begin{aligned} \sum_{i=1}^{2} F^i_{\mu\nu}F^{i\mu\nu} = 4\{&\partial_\mu W^+_\nu \partial^\mu W^{-\nu} - \partial_\mu W^+_\nu \partial^\nu W^{-\mu} - \mathrm{i}\left[e(\partial_\mu W^+_\nu W^{-\mu} A^\nu\right. \\ &- \partial_\mu W^+_\nu W^{-\nu} A^\mu) + g\cos\theta(\partial_\mu W^+_\nu W^{-\mu} Z^\nu - \partial_\mu W^+_\nu W^{-\nu} Z^\mu)] \\ &+ \mathrm{i}\left[e(\partial_\mu W^-_\nu W^{+\mu} A^\nu - \partial_\mu W^-_\nu W^{+\nu} A^\mu) + g\cos\theta(\partial_\mu W^-_\nu W^{+\mu} Z^\nu\right. \\ &\left. - \partial_\mu W^-_\nu W^{+\nu} Z^\mu)\right] + \left[e^2(W^+_\mu A_\nu W^{-\mu} A^\nu - W^+_\mu A_\nu W^{-\nu} A^\mu)\right. \\ &+ eg\cos\theta(2W^+_\mu W^{-\mu} A_\nu Z^\nu - W^+_\mu A_\nu W^{-\nu} Z^\mu - W^-_\mu A_\nu W^{+\nu} Z^\mu) \\ &+ g^2\cos^2\theta(W^+_\mu Z_\nu W^{-\mu} Z^\nu - W^+_\mu Z_\nu W^{-\nu} Z^\mu)]\} \end{aligned} \tag{8.7.9}$$

Using Eqs. (8.7.5c) and (8.7.5d), we can get

$$\begin{aligned}F^3_{\mu\nu}F^{3\mu\nu} &= 2\sin^2\theta\left(\partial_\mu A_\nu\partial^\mu A^\nu - \partial_\mu A_\nu\partial^\nu A^\mu\right)\\ &\quad + 2\cos^2\theta\left(\partial_\mu Z_\nu\partial^\mu Z^\nu - \partial_\mu Z_\nu\partial^\nu Z^\mu\right)\\ &\quad - 2g^2\left(W^+_\mu W^-_\nu W^{+\mu}W^{-\nu} - W^+_\mu W^-_\nu W^{-\mu}W^{+\nu}\right)\\ &\quad + 2\sin\theta\cos\theta\left(\partial_\mu A_\nu - \partial_\nu A_\mu\right)\left(\partial^\mu Z^\nu - \partial^\nu Z^\mu\right)\\ &\quad - 4\mathrm{i}e\left(\partial_\mu A_\nu W^{+\mu}W^{-\nu} - \partial_\mu A_\nu W^{-\mu}W^{+\nu}\right)\\ &\quad - 4\mathrm{i}g\cos\theta\left(\partial_\mu Z_\nu W^{+\mu}W^{-\nu} - \partial_\mu Z_\nu W^{-\mu}W^{+\nu}\right) \qquad (8.7.10\text{a})\end{aligned}$$

$$\begin{aligned}B_{\mu\nu}B^{\mu\nu} &= 2\cos^2\theta\left(\partial_\mu A_\nu\partial^\mu A^\nu - \partial_\mu A_\nu\partial^\nu A^\mu\right)\\ &\quad + 2\sin^2\theta\left(\partial_\mu Z_\nu\partial^\mu Z^\nu - \partial_\mu Z_\nu\partial^\nu Z^\mu\right)\\ &\quad - 2\sin\theta\cos\theta\left(\partial_\mu A_\nu - \partial_\nu A_\mu\right)\left(\partial^\mu Z^\nu - \partial^\nu Z^\mu\right) \qquad (8.7.10\text{b})\end{aligned}$$

With the above two equations added, the mixing terms of field A and Z will cancel, and one can obtain

$$\begin{aligned}&F^3_{\mu\nu}F^{3\mu\nu} + B_{\mu\nu}B^{\mu\nu}\\ &\quad = 2\left(\partial_\mu A_\nu\partial^\mu A^\nu - \partial_\mu A_\nu\partial^\nu A^\mu\right) + 2\left(\partial_\mu Z_\nu\partial^\mu Z^\nu - \partial_\mu Z_\nu\partial^\nu Z^\mu\right)\\ &\qquad - 2g^2\left(W^+_\mu W^{+\mu}W^-_\nu W^{-\nu} - W^+_\mu W^{-\mu}W^-_\nu W^{+\nu}\right)\\ &\qquad - 4\mathrm{i}e\left(\partial_\mu A_\nu W^{+\mu}W^{-\nu} - \partial_\mu A_\nu W^{-\mu}W^{+\nu}\right)\\ &\qquad - 4\mathrm{i}g\cos\theta\left(\partial_\mu Z_\nu W^{+\mu}W^{-\nu} - \partial_\mu Z_\nu W^{-\mu}W^{+\nu}\right) \qquad (8.7.11)\end{aligned}$$

The Lagrangian of pure gauge fields can be obtained by summing Eqs. (8.7.9) and (8.7.11)

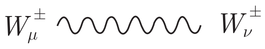
$$\begin{aligned}\mathcal{L}_{\mathrm{G}} = \partial_\mu W^+_\nu\partial^\nu W^{-\mu} - \partial_\mu W^+_\nu\partial^\mu W^{-\nu} + \frac{1}{2}(\partial_\mu A_\nu\partial^\nu A^\mu - \partial_\mu A_\nu\partial^\mu A^\nu)\\ + \frac{1}{2}(\partial_\mu Z_\nu\partial^\nu Z^\mu - \partial_\mu Z_\nu\partial^\mu Z^\nu) + \mathrm{i}e(\partial_\mu W^+_\nu W^{-\mu}A^\nu - \partial_\mu W^+_\nu W^{-\nu}A^\mu)\\ - \mathrm{i}e(\partial_\mu W^-_\nu W^{+\mu}A^\nu - \partial_\mu W^-_\nu W^{+\nu}A^\mu) + \mathrm{i}e(\partial_\mu A_\nu W^{+\mu}W^{-\nu}\\ - \partial_\mu A_\nu W^{-\mu}W^{+\nu}) + \mathrm{i}g\cos\theta(\partial_\mu W^+_\nu W^{-\mu}Z^\nu - \partial_\mu W^+_\nu W^{-\nu}Z^\mu)\\ - \mathrm{i}g\cos\theta(\partial_\mu W^-_\nu W^{+\mu}Z^\nu - \partial_\mu W^-_\nu W^{+\nu}Z^\mu)\end{aligned}$$

$$\begin{aligned}
&+ \mathrm{i}g\cos\theta\left(\partial_\mu Z_\nu W^{+\mu}W^{-\nu} - \partial_\mu Z_\nu W^{-\mu}W^{+\nu}\right)\\
&- e^2(W^+_\mu W^{-\mu}A_\nu A^\nu - W^+_\mu A_\nu W^{-\nu}A^\mu) - g^2\cos^2\theta(W^+_\mu Z_\nu W^{-\mu}Z^\nu\\
&- W^+_\mu Z_\nu W^{-\nu}Z^\mu) - eg\cos\theta(2W^+_\mu W^{-\mu}A_\nu Z^\nu - W^+_\mu A_\nu W^{-\nu}Z^\mu\\
&- W^-_\mu A_\nu W^{+\nu}Z^\mu) + \frac{1}{2}g^2\left(W^+_\mu W^{+\mu}W^-_\nu W^{-\nu} - W^+_\mu W^{-\mu}W^-_\nu W^{+\nu}\right)
\end{aligned} \tag{8.7.12}$$

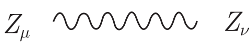
Equation (8.7.12) describes the interaction between gauge fields. The first three terms are the kinetic energy terms of $W^\pm$, photon and Z, respectively. The sum of the kinetic energy terms in Eq. (8.7.12) and the mass terms in Eq. (8.4.35) is

$$\begin{aligned}
&\partial_\mu W^+_\nu\partial^\nu W^{-\mu} - \partial_\mu W^+_\nu\partial^\mu W^{-\nu} + m_W^2 W^+_\mu W^{-\mu}\\
&\quad + \frac{1}{2}(\partial_\mu A_\nu\partial^\nu A^\mu - \partial_\mu A_\nu\partial^\mu A^\nu)\\
&\quad + \frac{1}{2}(\partial_\mu Z_\nu\partial^\nu Z^\mu - \partial_\mu Z_\nu\partial^\mu Z^\nu) + \frac{1}{2}m_Z^2 Z_\mu Z^\mu
\end{aligned} \tag{8.7.13}$$

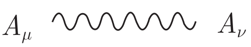
Equation (8.7.13) can give the propagators of gauge fields in the perturbation theory. In unitary gauge, these propagators are



$$\left(g_{\mu\nu} - \frac{k_\mu k_\nu}{m_W^2}\right)\frac{-\mathrm{i}}{k^2 - m_W^2 + \mathrm{i}\varepsilon}$$



$$\left(g_{\mu\nu} - \frac{k_\mu k_\nu}{m_Z^2}\right)\frac{-\mathrm{i}}{k^2 - m_Z^2 + \mathrm{i}\varepsilon}$$

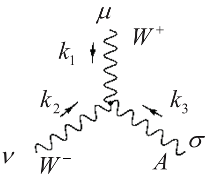


$$\left(g_{\mu\nu} - (1-\xi)\frac{k_\mu k_\nu}{k^2 + \mathrm{i}\varepsilon}\right)\frac{-\mathrm{i}}{k^2 + \mathrm{i}\varepsilon}$$

The parameter ξ in the propagator of photon is the gauge parameter, which can be set to arbitrary values. The different values for ξ will not lead to different physical results because of the gauge symmetry of the interaction.

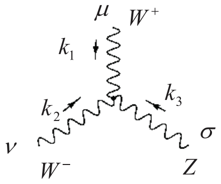
The other terms of Eq. (8.7.12) describe the interaction between the electroweak gauge fields, which give the Feynman rules for the interaction vertices between gauge fields:

①WWA vertex


μ
W^+
k_1
k_2
k_3
ν
W^-
A
σ

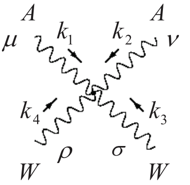
$$-ie[g^{\mu\nu}(k_2-k_1)^\sigma+g^{\nu\sigma}(k_3-k_2)^\mu+g^{\sigma\mu}(k_1-k_3)^\nu]$$

②WWZ vertex


μ
W^+
k_1
k_2
k_3
ν
σ
W^-
Z

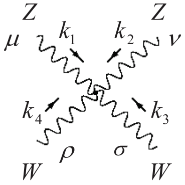
$$-ig\cos\theta\ [g^{\mu\nu}(k_2-k_1)^\sigma+g^{\nu\sigma}(k_3-k_2)^\mu+g^{\sigma\mu}(k_1-k_3)^\nu]$$

③WWAA vertex


A
A
μ
k_1
k_2
ν
k_4
k_3
ρ
σ
W
W

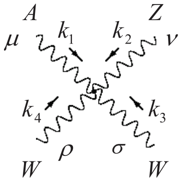
$$ie^2(g^{\mu\sigma}g^{\nu\rho}+g^{\mu\rho}g^{\nu\sigma}-2g^{\mu\nu}g^{\sigma\rho})$$

④WWZZ vertex


Z
Z
μ
k_1
k_2
ν
k_4
k_3
ρ
σ
W
W

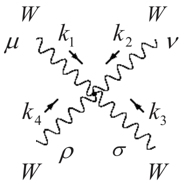
$$ig^2\cos^2\theta\ (g^{\mu\sigma}g^{\nu\rho}+g^{\mu\rho}g^{\nu\sigma}-2g^{\mu\nu}g^{\sigma\rho})$$

⑤WWAZ vertex


A
Z
μ
k_1
k_2
ν
k_4
k_3
ρ
σ
W
W

$$-ieg\cos\theta\ (2g^{\mu\nu}g^{\sigma\rho}-g^{\mu\sigma}g^{\nu\rho}-g^{\mu\rho}g^{\nu\sigma})$$

⑥WWWW vertex


W
W
μ
k_1
k_2
ν
k_4
k_3
ρ
σ
W
W

$$ig^2(2g^{\mu\nu}g^{\sigma\rho}-g^{\mu\sigma}g^{\nu\rho}-g^{\mu\rho}g^{\nu\sigma})$$

8.8. Feynman Rule in the Renormalization Gauge (R_ξ Gauge)

We have discussed the electroweak unification model in the unitary gauge and given the relevant Feynman rule in previous sections. In the unitary gauge, all the particles appearing in the theory are physical particles, and there are no non-physical particles in the theory. Non-physical particles, such as $\phi^\pm$ and χ, disappear, which makes the number of Feynman diagrams for the amplitude of any interaction processes greatly decrease. However, since there are terms like $-\frac{k_\mu k_\nu}{m_{\rm V}^2}$ in the propagators of vector bosons in the unitary gauge (see the propagators of $\rm W^\pm$ and $\rm Z^0$, and $m_{\rm V}$ is the mass of the vector boson), the vector's propagator tends to $O(1)$ rather than the asymptotic form $1/k^2$ in the limit of $k \to \infty$, which makes the degree of divergence of loop diagram increase when calculating the electroweak higher-order corrections. Therefore, it is inconvenient to calculate loop diagrams and discuss the problem of renormalization in the unitary gauge. In this section, we present the Feynman rules in non-unitary gauge, i.e. retaining the Goldstone particles $\phi^\pm$ and χ in the theory. Then the vertices and propagators involving Goldstone particles should be included in the Feynman rule. The gauge discussed in this section is referred to as renormalization gauge, or R_ξ gauge. In R_ξ gauge, the propagators of vector bosons will change, which makes the degree of divergence of the electroweak loop diagram greatly decrease. So R_ξ gauge is more appropriate to be used to calculate electroweak loop diagrams and discuss the problems of renormalization. In the meantime, there are gauge parameters apparently existing in the Feynman amplitudes in R_ξ gauge. But all the physical amplitudes should be gauge independent because of gauge symmetry. Therefore, the dependence of Feynman amplitudes on gauge parameters should eventually cancel each other, which provides a strong test for the validity of theoretical calculation.

After spontaneous symmetry breaking, if the scalar field is not taken as the form in Eq. (8.4.24), but as Eq. (8.4.10), i.e.

$$\Phi(x) \equiv \begin{pmatrix} \phi^+(x) \\ [\upsilon + H(x) + {\rm i}\chi(x)]/\sqrt{2} \end{pmatrix} \tag{8.8.1}$$

then Feynman rules that involve scalar field will change.

8.8.1. *For pure scalar field*

It can be seen from Eq. (8.4.20) that, compared with Eq. (8.4.25), there are more terms involving particles $\phi^{\pm}$ and χ in the Lagrangian for pure scalar fields in R_ξ gauge, which are

$$\begin{aligned}\mathcal{L}'_{\mathcal{H}} &= \partial_\mu\phi^-\partial^\mu\phi^+ + \frac{1}{2}\partial_\mu\chi\partial^\mu\chi - 2\upsilon\lambda\left(\phi^+\phi^- + \frac{1}{2}\chi^2\right)H \\ &\quad - \lambda\left[\left(\phi^+\phi^- + \frac{1}{2}\chi^2\right)^2 + \left(\phi^+\phi^- + \frac{1}{2}\chi^2\right)H^2\right] \\ &= \partial_\mu\phi^-\partial^\mu\phi^+ + \frac{1}{2}\partial_\mu\chi\partial^\mu\chi - \frac{m_H^2}{\upsilon}\left(\phi^+\phi^- + \frac{1}{2}\chi^2\right)H - \frac{m_H^2}{2\upsilon^2}\bigg[(\phi^+\phi^-)^2 \\ &\quad + \phi^+\phi^-\chi^2 + \frac{1}{4}\chi^4 + \phi^+\phi^-H^2 + \frac{1}{2}\chi^2H^2\bigg] \qquad (8.8.2)\end{aligned}$$

8.8.2. *For the interaction of scalar and gauge fields*

Substitute Eq. (8.8.1) into the covariant derivative term of the scalar field, then Eq. (8.4.41) becomes

$$\begin{aligned}&(D_\mu\Phi)^+D^\mu\Phi \\ &\quad = \left|\left[\partial_\mu - \mathrm{i}\begin{pmatrix} eA_\mu + g\dfrac{\cos 2\theta}{2\cos\theta}Z_\mu & \dfrac{g}{\sqrt{2}}W_\mu^+ \\ \dfrac{g}{\sqrt{2}}W_\mu^- & -\dfrac{g}{2\cos\theta}Z_\mu\end{pmatrix}\right]\begin{pmatrix}\phi^+ \\ \dfrac{\upsilon + H + \mathrm{i}\chi}{\sqrt{2}}\end{pmatrix}\right|^2 \\ &\quad = \left|\begin{pmatrix}\partial_\mu\phi^+ \\ \dfrac{\partial_\mu H + \mathrm{i}\partial_\mu\chi}{\sqrt{2}}\end{pmatrix} - \mathrm{i}\begin{pmatrix} eA_\mu\phi^+ + g\dfrac{\cos 2\theta}{2\cos\theta}Z_\mu\phi^+ + \dfrac{g}{\sqrt{2}} \\ W_\mu^+\dfrac{\upsilon + H + \mathrm{i}\chi}{\sqrt{2}} \\ \dfrac{g}{\sqrt{2}}W_\mu^-\phi^+ - \dfrac{g}{2\cos\theta}Z_\mu\dfrac{\upsilon + H + \mathrm{i}\chi}{\sqrt{2}}\end{pmatrix}\right|^2 \\ &\quad = \left|\begin{pmatrix}\partial_\mu\phi^+ - \mathrm{i}eA_\mu\phi^+ - \mathrm{i}g\dfrac{\cos 2\theta}{2\cos\theta}Z_\mu\phi^+ - \mathrm{i}\dfrac{g}{2}W_\mu^+(\upsilon + H + \mathrm{i}\chi) \\ \dfrac{1}{\sqrt{2}}\partial_\mu H + \dfrac{\mathrm{i}}{\sqrt{2}}\partial_\mu\chi - \mathrm{i}\dfrac{g}{\sqrt{2}}W_\mu^-\phi^+ + \mathrm{i}\dfrac{g}{2\sqrt{2}\cos\theta}Z_\mu(\upsilon + H + \mathrm{i}\chi)\end{pmatrix}\right|^2 \end{aligned}$$

$$(8.8.3)$$

Expanding Eq. (8.8.3) and simplifying it, we can obtain

$$
\begin{aligned}
(D_\mu\Phi)^+ D^\mu\Phi ={}& \frac{1}{2}\partial_\mu H\partial^\mu H + \partial_\mu\phi^+\partial^\mu\phi^- + \frac{1}{2}\partial_\mu\chi\partial^\mu\chi \\
&\left\{\begin{aligned}
&+\mathrm{i}\left(eA_\mu + g\frac{\cos 2\theta}{2\cos\theta}Z_\mu\right)(\partial^\mu\phi^+\phi^- - \partial^\mu\phi^-\phi^+) \\
&\quad+\frac{g}{2}W^{+\mu}(\partial_\mu\phi^-\chi - \phi^-\partial_\mu\chi) \\
&+\frac{g}{2}W^{-\mu}(\partial_\mu\phi^+\chi - \phi^+\partial_\mu\chi) + \mathrm{i}\frac{g}{2}W^{+\mu}(\partial_\mu H\phi^- - H\partial_\mu\phi^-) \\
&+\mathrm{i}\frac{g}{2}W^{-\mu}(H\partial_\mu\phi^+ - \partial_\mu H\phi^+) + \frac{g}{2\cos\theta}Z^\mu(\partial_\mu\chi H - \chi\partial_\mu H)
\end{aligned}\right. \\
&\left\{\begin{aligned}
&+\left(\frac{1}{2}eg\upsilon A_\mu - gm_Z\sin^2\theta Z_\mu\right)(W^{+\mu}\phi^- + W^{-\mu}\phi^+) \\
&+gm_W W^+_\mu W^{-\mu}H + \frac{g}{2\cos\theta}m_Z Z_\mu Z^\mu H
\end{aligned}\right. \\
&\left\{\begin{aligned}
&+\mathrm{i}\left(\frac{1}{2}eg\chi A_\mu - \frac{g^2\sin^2\theta}{2\cos\theta}\chi Z_\mu\right)(W^{+\mu}\phi^- - W^{-\mu}\phi^+) \\
&+\Bigg(e^2A_\mu A^\mu + eg\frac{\cos 2\theta}{\cos\theta}A_\mu Z^\mu + \left(\frac{g\cos 2\theta}{2\cos\theta}\right)^2 Z_\mu Z^\mu \\
&\quad+\frac{g^2}{2}W^+_\mu W^{-\mu}\Bigg)\phi^+\phi^- \\
&+\left(\frac{1}{2}egA_\mu H - \frac{g^2\sin^2\theta}{2\cos\theta}Z_\mu H\right)(W^{+\mu}\phi^- + W^{-\mu}\phi^+) \\
&+\frac{g^2}{4}W^+_\mu W^{-\mu}\chi^2 + \frac{g^2}{8\cos^2\theta}Z_\mu Z^\mu\chi^2 \\
&+\frac{g^2}{4}W^+_\mu W^{-\mu}H^2 + \frac{g^2}{8\cos^2\theta}Z_\mu Z^\mu H^2
\end{aligned}\right. \\
&+\mathrm{i}\frac{g\upsilon}{2}(\partial_\mu\phi^+W^{-\mu} - \partial_\mu\phi^-W^{+\mu}) + \frac{g\upsilon}{2\cos\theta}\partial_\mu\chi Z^\mu \\
&+\frac{g^2\upsilon^2}{4}W^+_\mu W^{-\mu} + \frac{g^2\upsilon^2}{8\cos^2\theta}Z_\mu Z^\mu
\end{aligned}
\tag{8.8.4}
$$

These three terms in the first line of Eq. (8.8.4) are the kinetic energy terms of Higgs, Goldstone particle $\phi^\pm$ and χ, respectively. The last line is for the mass terms of gauge fields $W^\pm$ and Z, and $m_W = \frac{1}{2}g\upsilon$, $m_Z = \frac{g\upsilon}{2\cos\theta}$. The second to third-to-last lines are terms for the interaction of gauge field with scalar particles. Compared with Eq. (8.4.42), more terms involving $\phi^\pm$ and

χ appear. The terms in the second-to-last line are the mixings of gauge field and scalar field. There is no direct physical interpretation for these terms. This reflects that in the case of containing non-physical particles $\phi^\pm$ and χ, there are non-physical degrees of freedom in the gauge field $W^\pm$ and Z, which can be eliminated by selecting appropriate gauge fixing terms.

The choice of the gauge fixing term is arbitrary in principle. In order to eliminate the quadratic mixing terms of the gauge field with $\phi^\pm$ and χ in Eq. (8.8.4)

$$\begin{aligned}&\mathrm{i}\frac{gv}{2}(\partial_\mu\phi^+W^{-\mu}-\partial_\mu\phi^-W^{+\mu})+\frac{gv}{2\cos\theta}\partial_\mu\chi Z^\mu\\&\quad=\mathrm{i}m_W(\partial_\mu\phi^+W^{-\mu}-\partial_\mu\phi^-W^{+\mu})+m_Z\partial_\mu\chi Z^\mu\end{aligned}\tag{8.8.5}$$

we choose the gauge fixing term as (how the gauge fixing term is introduced can be found in Chapter 9 on the discussion about quantization of gauge fields)

$$\mathcal{L}_{\text{fix}}=-\frac{1}{2}(F_\gamma^2+F_Z^2+2F_+F_-)\tag{8.8.6}$$

where

$$\begin{aligned}F_\gamma&=\frac{1}{\sqrt{\xi_\gamma}}G_\gamma, &G_\gamma&=\partial^\mu A_\mu\\F_Z&=\frac{1}{\sqrt{\xi_Z}}G_Z, &G_Z&=\partial^\mu Z_\mu-m_Z\xi_Z\chi\\F_\pm&=\frac{1}{\sqrt{\xi_W}}G_\pm, &G_\pm&=\partial^\mu W^\pm\mp\mathrm{i}m_W\xi_W\phi^\pm\end{aligned}\tag{8.8.7}$$

After eliminating the quadratic mixing term of the fields, the gauge fixing term becomes

$$\begin{aligned}\mathcal{L}_{\text{fix}}=&-\frac{1}{2\xi_\gamma}(\partial^\mu A_\mu)^2-\frac{1}{2\xi_Z}(\partial^\mu Z_\mu)^2-\frac{1}{\xi_W}(\partial^\mu W_\mu^+)(\partial^\nu W_\nu^-)\\&-\frac{1}{2}\xi_Z m_Z^2\chi^2-\xi_W m_W^2\phi^+\phi^-\end{aligned}\tag{8.8.8}$$

The last two terms in Eq. (8.8.8) are formally the mass terms of χ and $\phi^\pm$ respectively, but they all depend on the gauge fixing parameters $\xi_{Z,W}$. However, the gauge fixing parameter can be arbitrary, so this further indicates that χ and $\phi^\pm$ are non-physical particles, which cannot be generated in any interaction process.

8.8.3. *For ghost fields*

According to Faddeev and Popov's method [13], ghost fields should be introduced in the theory when a certain gauge is chosen for gauge fields. (The details of introducing ghost fields can be found in Chapter 9 about the discussion of the quantization of gauge field, or refer to Ref. [14], here only the main steps are presented.) The contribution of gauge fields to the Lagrangian is

$$\mathcal{L}_{gh} = -\bar{c}^a M^{ab} c^b \tag{8.8.9}$$

where the indices $a, b = \gamma, Z, +, -$, which are relevant to electromagnetic field, Z boson field and $W^{\pm}$ field, respectively. And c^b and $\bar{c}^a$ are ghost field and anti-ghost field. Although they are Lorentz scalars, they are anti-commutative complex numbers simultaneously. Note that c and $\bar{c}$ are not conjugates to each other. The matrix composed of M^{ab} is referred to as Faddeev–Popov matrix. The matrix element M^{ab} can be obtained by the derivative of the gauge fixing term with respect to gauge transformation parameters

$$M^{ab} = \frac{\partial G_a}{\partial \beta^b} \tag{8.8.10}$$

where β^b $(b = \gamma, Z, +, -)$ are relevant to the group parameters of U(1) and SU(2) gauge transformation, respectively. To derive the Faddeev–Popov matrix, one should know the variation $\mathrm{d}G_a$ of the gauge fixing function G_a in Eq. (8.8.7) under the infinitesimal gauge transformation. Then one needs to derive the variation of the scalar field $\Phi(x)$ and gauge field W_μ^a and B_μ under infinitesimal transformation at first.

The gauge transformation of scalar field is

$$\Phi(x) \to \Phi'(x) = \mathrm{e}^{\mathrm{i}gT^j\beta^j + \mathrm{i}\frac{g'}{2}\alpha}\Phi(x) \tag{8.8.11}$$

so under the infinitesimal transformation

$$\begin{aligned} \mathrm{d}\Phi &= \mathrm{i}\left(gT^j\beta^j + \frac{g'}{2}\alpha\right)\Phi(x) \\ &= \frac{\mathrm{i}}{2}\begin{pmatrix} g\beta^3 + g'\alpha & g(\beta^1 - \mathrm{i}\beta^2) \\ g(\beta^1 + \mathrm{i}\beta^2) & -g\beta^3 + g'\alpha \end{pmatrix}\Phi(x) \end{aligned} \tag{8.8.12}$$

Define

$$\begin{cases} \beta^{\pm} = \dfrac{1}{\sqrt{2}}(\beta^1 \mp \mathrm{i}\beta^2) \\ \beta_Z = -\alpha\sin\theta + \beta^3\cos\theta \\ \beta_\gamma = \alpha\cos\theta + \beta^3\sin\theta \end{cases} \tag{8.8.13}$$

then

$$\mathrm{d}\Phi = \mathrm{i}\begin{pmatrix} e'\beta_Z + e\beta_\gamma & \dfrac{g}{\sqrt{2}}\beta^+ \\ \dfrac{g}{\sqrt{2}}\beta^- & -\dfrac{g}{2\cos\theta}\beta_Z \end{pmatrix}\begin{pmatrix} \phi^+ \\ \dfrac{1}{\sqrt{2}}(\upsilon + H + \mathrm{i}\chi) \end{pmatrix}$$

$$= \mathrm{i}\begin{pmatrix} (e'\beta_Z + e\beta_\gamma)\phi^+ + \dfrac{1}{2}g\beta^+(\upsilon + H + \mathrm{i}\chi) \\ \dfrac{g}{\sqrt{2}}\beta^-\phi^+ - \dfrac{g}{2\sqrt{2}\cos\theta}\beta_Z(\upsilon + H + \mathrm{i}\chi) \end{pmatrix} \tag{8.8.14}$$

where

$$e' = \frac{g^2 - g'^2}{2\sqrt{g^2 + g'^2}} = \frac{g\cos 2\theta}{2\cos\theta}, \quad e = \frac{gg'}{\sqrt{g^2 + g'^2}}$$

Since

$$\mathrm{d}\Phi \equiv \begin{pmatrix} \mathrm{d}\phi^+ \\ \dfrac{1}{\sqrt{2}}(\mathrm{d}H + \mathrm{i}\mathrm{d}\chi) \end{pmatrix} \tag{8.8.15}$$

then, comparing Eqs. (8.8.15) and (8.8.14), we can obtain

$$\begin{cases} \mathrm{d}\phi^+ = \mathrm{i}(e'\beta_Z + e\beta_\gamma)\phi^+ + \dfrac{\mathrm{i}}{2}g\beta^+(\upsilon + H + \mathrm{i}\chi) \\ \mathrm{d}\phi^- = -\mathrm{i}(e'\beta_Z + e\beta_\gamma)\phi^- - \dfrac{\mathrm{i}}{2}g\beta^-(\upsilon + H - \mathrm{i}\chi) \\ \mathrm{d}\chi = \dfrac{1}{2}g(\beta^-\phi^+ + \beta^+\phi^-) - \dfrac{g}{2\cos\theta}\beta_Z(\upsilon + H) \end{cases} \tag{8.8.16}$$

The infinitesimal transformation of the gauge fields can be derived from Eqs. (8.3.15) and (8.3.16). From Eq. (8.3.15), we can get

$$\mathrm{d}B_\mu = \partial_\mu\alpha(x) \tag{8.8.17}$$

Substituting the SU(2) infinitesimal transformation $U = 1 + \mathrm{i}gT^j\beta^j(x)$ into Eq. (8.3.16) and simplifying it, we can obtain

$$T^jA'^j_\mu = T^j(A^j_\mu + g\varepsilon^{lkj}A^l_\mu\beta^k + \partial_\mu\beta^j)$$

then we get

$$\mathrm{d}A^j_\mu = g\varepsilon^{jkl}A^k_\mu\beta^l + \partial_\mu\beta^j \tag{8.8.18}$$

It can be derived from Eq. (8.8.18)

$$\mathrm{d}W^+_\mu = \frac{1}{\sqrt{2}}(\mathrm{d}A^1_\mu - \mathrm{i}\mathrm{d}A^2_\mu) = \partial_\mu\beta^+ + \mathrm{i}gW^+_\mu\beta^3 - \mathrm{i}gA^3_\mu\beta^+ \tag{8.8.19a}$$

$$\mathrm{d}W^-_\mu = \frac{1}{\sqrt{2}}(\mathrm{d}A^1_\mu + \mathrm{i}\mathrm{d}A^2_\mu) = \partial_\mu\beta^- - \mathrm{i}gW^-_\mu\beta^3 + \mathrm{i}gA^3_\mu\beta^- \tag{8.8.19b}$$

$$\mathrm{d}A^3_\mu = \partial_\mu\beta^3 + g(A^1_\mu\beta^2 - A^2_\mu\beta^1) = \partial_\mu\beta^3 + \mathrm{i}g(W^-_\mu\beta^+ - W^+_\mu\beta^-) \tag{8.8.20}$$

With Eqs. (8.3.18), (8.3.17) and (8.8.20), we can obtain the infinitesimal transformation of Z and electromagnetic field,

$$\begin{aligned}
\mathrm{d}Z_\mu &= -\sin\theta\mathrm{d}B_\mu + \cos\theta\mathrm{d}A^3_\mu = \partial_\mu\beta_Z + \mathrm{i}g\cos\theta(W^-_\mu\beta^+ - W^+_\mu\beta^-)\\
\mathrm{d}A_\mu &= \cos\theta\mathrm{d}B_\mu + \sin\theta\mathrm{d}A^3_\mu = \partial_\mu\beta_\gamma + \mathrm{i}e(W^-_\mu\beta^+ - W^+_\mu\beta^-)
\end{aligned} \tag{8.8.21}$$

Now let us find the change of the gauge fixing term G_a under the infinitesimal gauge transformation

$$\begin{aligned}
\mathrm{d}G_\gamma &= \mathrm{d}(\partial^\mu A_\mu) = \partial^\mu(\mathrm{d}A_\mu)\\
&= \partial^\mu\partial_\mu\beta_\gamma + \mathrm{i}e\partial^\mu(W^-_\mu\beta^+ - W^+_\mu\beta^-)\\
&= \partial^\mu\partial_\mu\beta_\gamma + \mathrm{i}e(\partial^\mu W^-_\mu\beta^+ + W^-_\mu\partial^\mu\beta^+) - \mathrm{i}e(\partial^\mu W^+_\mu\beta^- + W^+_\mu\partial^\mu\beta^-)\\
&= \partial^\mu\partial_\mu\beta_\gamma - \mathrm{i}e(\partial^\mu W^+_\mu + W^+_\mu\partial^\mu)\beta^- + \mathrm{i}e(\partial^\mu W^-_\mu + W^-_\mu\partial^\mu)\beta^+
\end{aligned} \tag{8.8.22a}$$

$$\begin{aligned}
\mathrm{d}G_Z &= \mathrm{d}(\partial^\mu Z_\mu - m_Z\xi_Z\chi) = \partial^\mu\mathrm{d}Z_\mu - m_Z\xi_Z\mathrm{d}\chi\\
&= \partial^\mu[\partial_\mu\beta_Z + \mathrm{i}g\cos\theta(W^-_\mu\beta^+ - W^+_\mu\beta^-)]\\
&\quad - m_Z\xi_Z\left[\frac{1}{2}g(\beta^-\phi^+ + \beta^+\phi^-) - \frac{g}{2\cos\theta}\beta_Z(\upsilon + H)\right]\\
&= \left(\partial^\mu\partial_\mu + \xi_Z m_Z^2 + \frac{g}{2\cos\theta}m_Z\xi_Z H\right)\beta_Z - \left[\mathrm{i}g\cos\theta(\partial^\mu W^+_\mu + W^+_\mu\partial^\mu)\right.\\
&\quad \left. + \frac{1}{2}gm_Z\xi_Z\phi^+\right]\beta^- + \left[\mathrm{i}g\cos\theta(\partial^\mu W^-_\mu + W^-_\mu\partial^\mu) - \frac{1}{2}gm_Z\xi_Z\phi^-\right]\beta^+
\end{aligned} \tag{8.8.22b}$$

The result $m_Z = \frac{1}{2}\upsilon\sqrt{g^2 + g'^2} = \frac{g\upsilon}{2\cos\theta}$ has been used in Eq. (8.8.22b).

$$\begin{aligned}
\mathrm{d}G_+ &= \mathrm{d}(\partial^\mu W^+ - \mathrm{i}m_W\xi_W\phi^+) = \partial^\mu\mathrm{d}W^+ - \mathrm{i}m_W\xi_W\mathrm{d}\phi^+\\
&= \partial^\mu[\partial_\mu\beta^+ + \mathrm{i}gW^+_\mu\beta^3 - \mathrm{i}gA^3_\mu\beta^+] - \mathrm{i}m_W\xi_W\left[\mathrm{i}(e'\beta_Z + e\beta_\gamma)\phi^+\right.\\
&\quad \left. + \frac{\mathrm{i}}{2}g\beta^+(\upsilon + H + \mathrm{i}\chi)\right]
\end{aligned}$$

$$= [\mathrm{i}e(\partial^\mu W_\mu^+ + W_\mu^+\partial^\mu) + em_W\xi_W\phi^+]\beta_\gamma + [\mathrm{i}g\cos\theta(\partial^\mu W_\mu^+ + W_\mu^+\partial^\mu)$$
$$+ e' m_W\xi_W\phi^+]\beta_Z + \left[\partial^\mu\partial_\mu - \mathrm{i}g(\partial^\mu A_\mu^3 + A_\mu^3\partial^\mu)\right.$$
$$\left. + \frac{1}{2}gm_W\xi_W(H+\mathrm{i}\chi)m_W^2\xi_W\right]\beta^+ \tag{8.8.22c}$$

In the derivation of Eq. (8.8.22c), we have used $m_W = \dfrac{1}{2}gv$ and

$$\beta^3 = \beta_Z\cos\theta + \beta_\gamma\sin\theta$$

which is derived from the last two equations of (8.8.13). Sine $\mathrm{d}G_-$and $\mathrm{d}G_+$ are conjugates of each other, so we have

$$\mathrm{d}G_- = (\mathrm{d}G_+)^+ \tag{8.8.22d}$$

According to the definition of the Faddeev–Popov matrix, (8.8.10) and the infinitesimal transformations obtained above, we can derive the Faddeev–Popov matrix. We define $M^{a\pm}$ as

$$M^{a\pm} = \frac{\partial G_a}{\partial\beta^\mp} \tag{8.8.23}$$

The final results of the matrix M^{ab} are

$$\begin{cases} M^{\gamma\gamma} = \partial^\mu\partial_\mu, \quad M^{\gamma Z} = 0 \\ M^{\gamma +} = -\mathrm{i}e(\partial^\mu W_\mu^+ + W_\mu^+\partial^\mu) \\ M^{\gamma -} = \mathrm{i}e(\partial^\mu W_\mu^- + W_\mu^-\partial^\mu) \end{cases} \tag{8.8.24}$$

$$\begin{cases} M^{Z\gamma} = 0, M^{ZZ} = \partial^\mu\partial_\mu + m_Z^2\xi_Z + \dfrac{\mathrm{g}}{2\mathrm{cos}\theta}m_Z\xi_Z H \\ M^{Z+} = -\mathrm{i}g\cos\theta(\partial^\mu W_\mu^+ + W_\mu^+\partial^\mu) - \dfrac{1}{2}gm_Z\xi_Z\phi^+ \\ M^{Z-} = \mathrm{i}g\cos\theta(\partial^\mu W_\mu^- + W_\mu^-\partial^\mu) - \dfrac{1}{2}gm_Z\xi_Z\phi^- \end{cases} \tag{8.8.25}$$

$$\begin{cases} M^{+\gamma} = \mathrm{i}e(\partial^\mu W_\mu^+ + W_\mu^+\partial^\mu) + em_W\xi_W\phi^+ \\ M^{+Z} = \mathrm{i}g\cos\theta(\partial^\mu W_\mu^+ + W_\mu^+\partial^\mu) + \dfrac{\mathrm{gcos}2\theta}{2\mathrm{cos}\theta}m_W\xi_W\phi^+ \\ M^{++} = 0 \\ M^{+-} = \partial^\mu\partial_\mu - \mathrm{i}g(\partial^\mu A_\mu^3 + A_\mu^3\partial^\mu) + \dfrac{1}{2}gm_W\xi_W(H+\mathrm{i}\chi) + m_W^2\xi_W \end{cases} \tag{8.8.26}$$

$$\begin{cases} M^{-\gamma} = -\mathrm{i}e(\partial^\mu W_\mu^- + W_\mu^-\partial^\mu) + em_W\xi_W\phi^- \\ M^{-Z} = -\mathrm{i}g\cos\theta(\partial^\mu W_\mu^- + W_\mu^-\partial^\mu) + \dfrac{g\cos 2\theta}{2\cos\theta}m_W\xi_W\phi^- \\ M^{--} = 0 \\ M^{-+} = \partial^\mu\partial_\mu + \mathrm{i}g(\partial^\mu A_\mu^3 + A_\mu^3\partial^\mu) + \dfrac{1}{2}gm_W\xi_W(H - \mathrm{i}\chi) + m_W^2\xi_W \end{cases} \tag{8.8.27}$$

Then the contribution of the ghost fields to the Lagrangian is

$$\begin{aligned} \mathcal{L}_{gh} = -\big[&\bar{c}^z M^{zz}c^z + \bar{c}^z M^{z+}c^- + \bar{c}^z M^{z-}c^+ + \bar{c}^\gamma M^{\gamma\gamma}c^\gamma \\ &+ \bar{c}^\gamma M^{\gamma+}c^- + \bar{c}^\gamma M^{\gamma-}c^+ + \bar{c}^- M^{+-}c^+ + \bar{c}^- M^{+z}c^z \\ &+ \bar{c}^- M^{+\gamma}c^\gamma + \bar{c}^+ M^{-+}c^- + \bar{c}^+ M^{-z}c^z + \bar{c}^+ M^{-\gamma}c^\gamma\big] \end{aligned} \tag{8.8.28}$$

8.8.4. *The kinetic energy term of the Lagrangian in R_ξ gauge*

The gauge fixing terms in Eq. (8.8.8) will contribute extra kinetic terms of vector and scalar fields, while Eq. (8.8.28) contributes to the ghost fields' kinetic energy and interaction terms. The kinetic energy terms of the vector and scalar fields in the Lagrangian can be summarized as follows:

$$\begin{aligned} &\partial_\mu W_\nu^+\partial^\nu W^{-\mu} - \partial_\mu W_\nu^+\partial^\mu W^{-\nu} + m_W^2 W_\mu^+ W^{-\mu} - \frac{1}{\xi_W}(\partial^\mu W_\mu^+)(\partial^\nu W_\nu^-) \\ &+ \frac{1}{2}(\partial_\mu Z_\nu\partial^\nu Z^\mu - \partial_\mu Z_\nu\partial^\mu Z^\nu) + \frac{1}{2}m_Z^2 Z_\mu Z^\mu - \frac{1}{2\xi_Z}(\partial^\mu Z_\mu)(\partial^\nu Z_\nu) \\ &+ \frac{1}{2}(\partial_\mu A_\nu\partial^\nu A^\mu - \partial_\mu A_\nu\partial^\mu A^\nu) - \frac{1}{2\xi_\gamma}(\partial^\mu A_\mu)^2 \\ &+ \partial_\mu\phi^+\partial^\mu\phi^- - m_W^2\xi_W\phi^+\phi^- + \frac{1}{2}\partial_\mu\chi\partial^\mu\chi - \frac{1}{2}m_Z^2\xi_Z\chi^2 \\ &+ \frac{1}{2}\partial_\mu H\partial^\mu H - \frac{1}{2}m_H^2 H^2 - \bar{c}^\gamma\partial^\mu\partial_\mu c^\gamma - \bar{c}^Z\partial^\mu\partial_\mu c^Z - m_Z^2\xi_Z\bar{c}^Z c^Z \\ &- \bar{c}^-\partial^\mu\partial_\mu c^+ - m_W^2\xi_W\bar{c}^- c^+ - \bar{c}^+\partial^\mu\partial_\mu c^- - m_W^2\xi_W\bar{c}^+ c^- \end{aligned} \tag{8.8.29}$$

while there is no change for the fermion's kinetic terms. The propagators of gauge fields $W^\pm$, Z and electromagnetic field A, and scalar and ghost fields can be derived from the kinetic terms.

8.8.5. *The interaction terms of scalar and fermions*

If Goldstone particles $\phi^{\pm}$ and χ are kept in the scalar doublet $\Phi(x)$, then, with spontaneous symmetry breaking

$$\Phi \equiv \begin{pmatrix} \phi^{+} \\ [\upsilon + H(x) + \mathrm{i}\chi]/\sqrt{2} \end{pmatrix} \tag{8.8.30}$$

According to Eq. (8.5.16), the conjugate scalar field $\tilde{\Phi}$ in Yukawa coupling in Eq. (8.5.15) becomes

$$\tilde{\Phi} \equiv \begin{pmatrix} [\upsilon + H(x) - \mathrm{i}\chi]/\sqrt{2} \\ -\phi^{-} \end{pmatrix} \tag{8.8.31}$$

Substituting Eqs. (8.8.30) and (8.8.31) into Eq. (8.5.15), we can get

$$\begin{aligned} \mathcal{L}_{\text{Yukawa}} = {} & -g_i^{\mathrm{L}} \bar{l}_{i\mathrm{R}} \left(\phi^{-}, \frac{\upsilon + H - \mathrm{i}\chi}{\sqrt{2}} \right) \begin{pmatrix} \nu_{i\mathrm{L}} \\ l_{i\mathrm{L}} \end{pmatrix} \\ & - g_{ij}^{d} \bar{d}'_{i\mathrm{R}} \left(\phi^{-}, \frac{\upsilon + H - \mathrm{i}\chi}{\sqrt{2}} \right) \begin{pmatrix} u_{j\mathrm{L}} \\ d'_{j\mathrm{L}} \end{pmatrix} \\ & - g_i^{u} \bar{u}_{i\mathrm{R}} \left(\frac{\upsilon + H + \mathrm{i}\chi}{\sqrt{2}}, -\phi^{+} \right) \begin{pmatrix} u_{i\mathrm{L}} \\ d'_{i\mathrm{L}} \end{pmatrix} + \text{h.c.} \end{aligned} \tag{8.8.32}$$

Compared with Eq. (8.5.19), there are more terms about the interaction of $\phi^{\pm}$ and χ. The extra new terms are

$$\begin{aligned} \mathcal{L}'_{\text{Yukawa}} = {} & -g_i^{\mathrm{L}} \bar{l}_{i\mathrm{R}} \left(\phi^{-}, \frac{-\mathrm{i}\chi}{\sqrt{2}} \right) \begin{pmatrix} \nu_{i\mathrm{L}} \\ l_{i\mathrm{L}} \end{pmatrix} - g_{ij}^{d} \bar{d}'_{i\mathrm{R}} \left(\phi^{-}, \frac{-\mathrm{i}\chi}{\sqrt{2}} \right) \begin{pmatrix} u_{j\mathrm{L}} \\ d'_{j\mathrm{L}} \end{pmatrix} \\ & - g_i^{u} \bar{u}_{i\mathrm{R}} \left(\frac{\mathrm{i}\chi}{\sqrt{2}}, -\phi^{+} \right) \begin{pmatrix} u_{i\mathrm{L}} \\ d'_{i\mathrm{L}} \end{pmatrix} + \text{h.c.} \\ = {} & -g_i^{\mathrm{L}} \bar{l}_{i\mathrm{R}} \nu_{i\mathrm{L}} \phi^{-} + \frac{\mathrm{i}}{\sqrt{2}} g_i^{\mathrm{L}} \bar{l}_{i\mathrm{R}} l_{i\mathrm{L}} \chi - g_{ij}^{d} \bar{d}'_{i\mathrm{R}} u_{j\mathrm{L}} \phi^{-} + \frac{i}{\sqrt{2}} g_{ij}^{d} \bar{d}'_{i\mathrm{R}} d'_{j\mathrm{L}} \chi \\ & - \frac{\mathrm{i}}{\sqrt{2}} g_i^{u} \bar{u}_{i\mathrm{R}} u_{i\mathrm{L}} \chi + g_i^{u} \bar{u}_{i\mathrm{R}} d'_{i\mathrm{L}} \phi^{+} + \text{h.c.} \end{aligned} \tag{8.8.33}$$

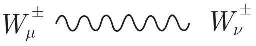
Expressing d' by its mass eigenstate, that is, considering Eq. (8.5.21), and using Eqs. (8.5.23) and (8.6.3), we can obtain

$$\begin{aligned}
\mathcal{L}'_{\text{Yukawa}} &= -\frac{\sqrt{2}m_{li}}{\upsilon}\bar{l}_{i\mathrm{R}}\nu_{i\mathrm{L}}\phi^- + \frac{\mathrm{i}m_{li}}{\upsilon}\bar{l}_{i\mathrm{R}}l_{i\mathrm{L}}\chi - \frac{\sqrt{2}m_{dj}}{\upsilon}V^*_{ij}\bar{d}_{j\mathrm{R}}u_{i\mathrm{L}}\phi^- \\
&\quad + \frac{\mathrm{i}m_{di}}{\upsilon}\bar{d}_{i\mathrm{R}}d_{i\mathrm{L}}\chi - \frac{\mathrm{i}m_{ui}}{\upsilon}\bar{u}_{i\mathrm{R}}u_{i\mathrm{L}}\chi + \frac{\sqrt{2}m_{ui}}{\upsilon}V_{ij}\bar{u}_{i\mathrm{R}}d_{j\mathrm{L}}\phi^+ + \text{h.c.} \\
&= -\frac{m_{li}}{\sqrt{2}\upsilon}\bar{\nu}_i(1+\gamma_5)l_i\phi^+ - \frac{m_{li}}{\sqrt{2}\upsilon}\bar{l}_i(1-\gamma_5)\nu_i\phi^- \\
&\quad + \frac{1}{\sqrt{2}\upsilon}V_{ij}\bar{u}_i[(m_{ui}-m_{dj}) - \gamma_5(m_{ui}+m_{dj})]d_j\phi^+ \\
&\quad + \frac{1}{\sqrt{2}\upsilon}V^*_{ij}\bar{d}_j[(m_{ui}-m_{dj}) + \gamma_5(m_{ui}+m_{dj})]u_i\phi^- \\
&\quad - \frac{\mathrm{i}}{\upsilon}[m_{li}\bar{l}_i\gamma_5 l_i + m_{di}\bar{d}_i\gamma_5 d_i - m_{ui}\bar{u}_i\gamma_5 u_i]\chi \qquad (8.8.34)
\end{aligned}$$

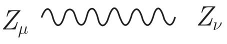
New terms about the interaction between fermion field and scalar fields $\phi^{\pm}$ and χ in R_ξ gauge are presented in Eq. (8.8.34).

The new interaction terms in R_ξ gauge where the scalar fields $\phi^{\pm}$ and χ are involved have been given above, while all the other terms remain the same as that in unitary gauge. The new terms contribute to new Feynman rules. Compared to unitary gauge, the propagators of vector fields change because of the introduction of the gauge fixing terms, while the fermions' propagators keep unchanged. We list the propagators and vertices that are changed and newly added as follows. The other Feynman rules are the same as that in unitary rules.

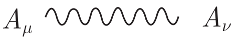
Propagator:



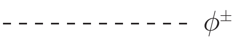
$$\left(g_{\mu\nu} - (1-\xi_W)\frac{k_\mu k_\nu}{k^2 - \xi_W m_W^2 + \mathrm{i}\varepsilon}\right)\frac{-\mathrm{i}}{k^2 - m_W^2 + \mathrm{i}\varepsilon}$$



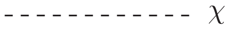
$$\left(g_{\mu\nu} - (1-\xi_Z)\frac{k_\mu k_\nu}{k^2 - \xi_Z m_Z^2 + \mathrm{i}\varepsilon}\right)\frac{-\mathrm{i}}{k^2 - m_Z^2 + \mathrm{i}\varepsilon}$$



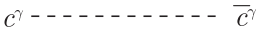
$$\left(g_{\mu\nu} - (1-\xi_\gamma)\frac{k_\mu k_\nu}{k^2 + \mathrm{i}\varepsilon}\right)\frac{-\mathrm{i}}{k^2 + \mathrm{i}\varepsilon}$$



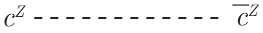
$$\frac{\mathrm{i}}{k^2 - \xi_W m_W^2 + \mathrm{i}\varepsilon}$$



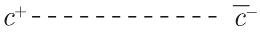
$$\frac{\mathrm{i}}{k^2-\xi_Z m_Z^2+\mathrm{i}\varepsilon}$$


c^γ
$\overline{c}^\gamma$

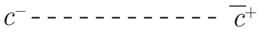
$$\frac{\mathrm{i}}{k^2+\mathrm{i}\varepsilon}$$


c^Z
$\overline{c}^Z$

$$\frac{\mathrm{i}}{k^2-\xi_Z m_Z^2+\mathrm{i}\varepsilon}$$


c^+
$\overline{c}^-$

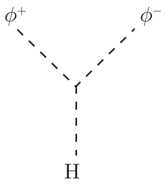
$$\frac{\mathrm{i}}{k^2-\xi_W m_W^2+\mathrm{i}\varepsilon}$$


c^-
$\overline{c}^+$

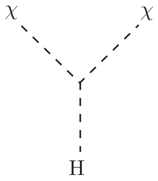
$$\frac{\mathrm{i}}{k^2-\xi_W m_W^2+\mathrm{i}\varepsilon}$$

The new added vertices:

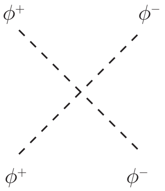
(1) *The interaction vertices of scalars* (*see Eq.* (8.8.2)):



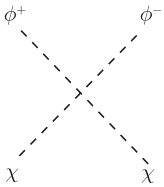
$$-\mathrm{i}\,\frac{m_{\mathrm{H}}^2}{v}$$



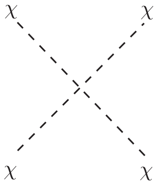
$$-\mathrm{i}\,\frac{m_{\mathrm{H}}^2}{v}$$


ϕ^+
ϕ^-
ϕ^+
ϕ^-

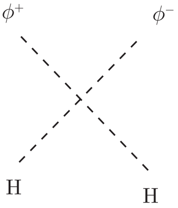
$$-2\mathrm{i}\,\frac{m_{\mathrm{H}}^2}{v^2}$$


ϕ^+
ϕ^-
χ
χ

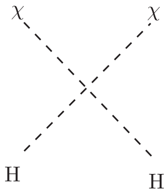
$$-\mathrm{i}\,\frac{m_{\mathrm{H}}^2}{v^2}$$


χ
χ
χ
χ

$$-3\mathrm{i}\,\frac{m_{\mathrm{H}}^2}{v^2}$$



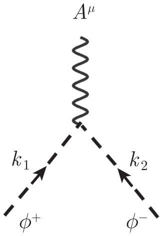
$$-\mathrm{i}\,\frac{m_{\mathrm{H}}^2}{v^2}$$



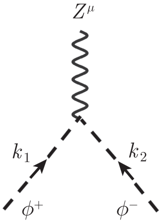
$$-\mathrm{i}\,\frac{m_{\mathrm{H}}^2}{v^2}$$

(2) *The vertices for the interaction of scalar and gauge fields*: According to Eq. (8.8.4), the new added vertices in R_ξ gauge include:

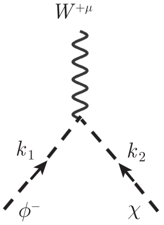
(i) *The vertices relevant to first-order differential terms*:


A^μ
k_1
k_2
ϕ^+
ϕ^-

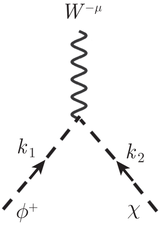
$-\mathrm{i}e(k_1^\mu - k_2^\mu)$


Z^μ
k_1
k_2
ϕ^+
ϕ^-

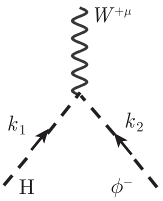
$-\mathrm{i}g\,\dfrac{\cos 2\theta}{2\cos\theta}(k_1^\mu - k_2^\mu)$


$W^{+\mu}$
k_1
k_2
ϕ^-
χ

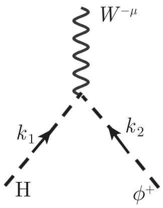
$-\dfrac{g}{2}(k_1^\mu - k_2^\mu)$


$W^{-\mu}$
k_1
k_2
ϕ^+
χ

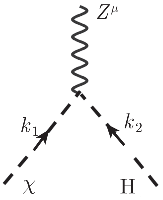
$-\dfrac{g}{2}(k_1^\mu - k_2^\mu)$


$W^{+\mu}$
k_1
k_2
H
ϕ^-

$-\dfrac{\mathrm{i}g}{2}\left(k_1^\mu - k_2^\mu\right)$

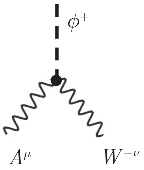

$W^{-\mu}$
k_1
k_2
H
ϕ^+

$\dfrac{\mathrm{i}g}{2}\left(k_1^\mu - k_2^\mu\right)$

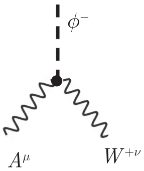

Z^μ
k_1
k_2
χ
H

$-\dfrac{g}{2\cos\theta}(k_1^\mu - k_2^\mu)$

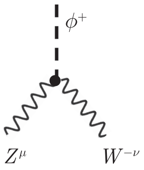
(ii) *The vertices relevant to cubic filed terms without derivative*:



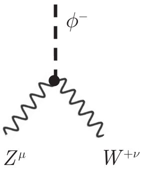
$\mathrm{i}em_W\,g^{\mu\nu}$



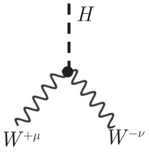
$\mathrm{i}em_W\,g^{\mu\nu}$



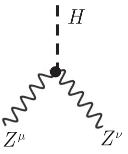
$-\mathrm{i}gm_Z\sin^2\theta g^{\mu\nu}$



$-\mathrm{i}gm_Z\sin^2\theta g^{\mu\nu}$

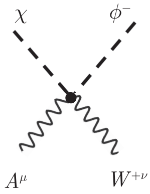


$\mathrm{i}gm_W\,g^{\mu\nu}$

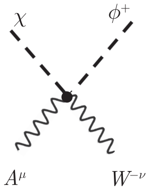


$\mathrm{i}\,\dfrac{g}{\cos\theta}\,m_Z\,g^{\mu\nu}$

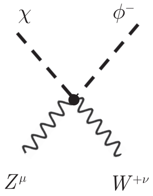
(iii) *The vertices relevant to quadratic filed terms without derivative*:


χ
ϕ^-
A^μ
$W^{+\nu}$

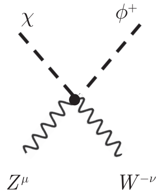
$$-\frac{1}{2}egg^{\mu\nu}$$


χ
ϕ^+
A^μ
$W^{-\nu}$

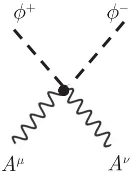
$$\frac{1}{2}egg^{\mu\nu}$$


χ
ϕ^-
Z^μ
$W^{+\nu}$

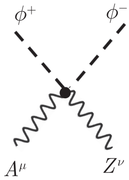
$$\frac{g^2\sin^2\theta}{2\cos\theta}g^{\mu\nu}$$


χ
ϕ^+
Z^μ
$W^{-\nu}$

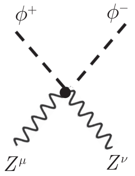
$$-\frac{g^2\sin^2\theta}{2\cos\theta}g^{\mu\nu}$$


ϕ^+
ϕ^-
A^μ
A^ν

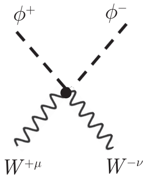
$$2\mathrm{i}e^2g^{\mu\nu}$$


ϕ^+
ϕ^-
A^μ
Z^ν

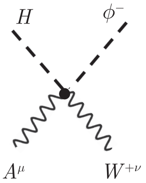
$$\mathrm{i}eg\frac{\cos2\theta}{\cos\theta}g^{\mu\nu}$$


ϕ^+
ϕ^-
Z^μ
Z^ν

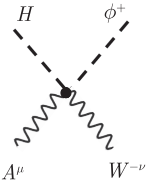
$$\mathrm{i}\frac{1}{2}\left(\frac{g\cos2\theta}{\cos\theta}\right)^2g^{\mu\nu}$$


ϕ^+
ϕ^-
$W^{+\mu}$
$W^{-\nu}$

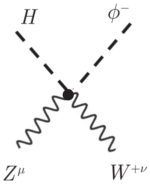
$$\mathrm{i}\frac{1}{2}g^2g^{\mu\nu}$$


H
ϕ^-
A^μ
$W^{+\nu}$

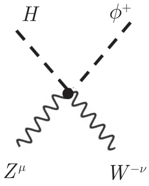
$$\mathrm{i}\frac{1}{2}egg^{\mu\nu}$$


H
ϕ^+
A^μ
$W^{-\nu}$

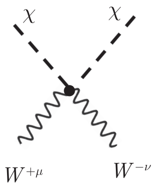
$$\mathrm{i}\frac{1}{2}egg^{\mu\nu}$$


H
ϕ^-
Z^μ
$W^{+\nu}$

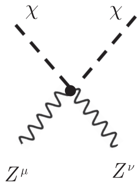
$$-\mathrm{i}\frac{g^2\sin^2\theta}{2\cos\theta}g^{\mu\nu}$$


H
ϕ^+
Z^μ
$W^{-\nu}$

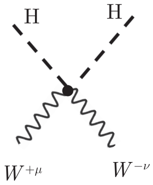
$$-\mathrm{i}\frac{g^2\sin^2\theta}{2\cos\theta}g^{\mu\nu}$$


χ
χ
$W^{+\mu}$
$W^{-\nu}$

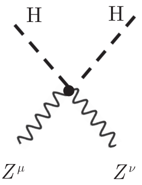
$$\mathrm{i}\frac{1}{2}g^2g^{\mu\nu}$$


χ
χ
Z^μ
Z^ν

$$\mathrm{i}\frac{g^2}{2\cos^2\theta}g^{\mu\nu}$$



$$\mathrm{i}\frac{1}{2}g^2g^{\mu\nu}$$



$$\mathrm{i}\frac{g}{2\cos^2\theta}g^{\mu\nu}$$

(3) *Vertices with ghost fields*: The Lagrangian of the ghost field is $\mathcal{L}_{gh} = -\bar{c}^aM^{ab}c^b$, where the Faddeev–Popov matrices are given by Eqs. (8.8.24)–(8.8.27). The kinetic terms of ghost fields have been separated out, and the other terms give the interaction vertices involving ghost fields. In the

interaction between ghost and gauge fields, the combination of

$$\bar{c}(\partial^\mu W_\mu^+ + W_\mu^+ \partial^\mu)c$$

may appear. Such term can be treated as follows:

$$\bar{c}(\partial^\mu W_\mu^+ + W_\mu^+ \partial^\mu)c = \bar{c}\partial^\mu(W_\mu^+ c) = \partial^\mu(\bar{c}W_\mu^+ c) - (\partial^\mu \bar{c})W_\mu^+ c$$

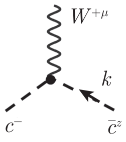
The contribution of the total differential in the above equation to the action is zero, so it can be removed from the Lagrangian, therefore

$$\bar{c}(\partial^\mu W_\mu^+ + W_\mu^+ \partial^\mu)c \to -(\partial^\mu \bar{c})W_\mu^+ c$$

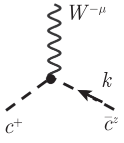
(i) *The interaction vertex of ghost and gauge fields*: According to Eqs. (8.8.24)–(8.8.28) and the above discussion, we can summarize the interaction between ghost and gauge fields as follows:

$$\begin{aligned}
&-\mathrm{i}g\cos\theta(\partial^\mu\bar{c}^z)W_\mu^+c^- + \mathrm{i}g\cos\theta(\partial^\mu\bar{c}^z)W_\mu^-c^+ - \mathrm{i}e(\partial^\mu\bar{c}^\gamma)W_\mu^+c^-\\
&+\mathrm{i}e(\partial^\mu\bar{c}^\gamma)W_\mu^-c^+ + \mathrm{i}e(\partial^\mu\bar{c}^-)W_\mu^+c^\gamma + \mathrm{i}g\cos\theta(\partial^\mu\bar{c}^-)W_\mu^+c^z\\
&-\mathrm{i}e(\partial^\mu\bar{c}^-)A_\mu c^+ - \mathrm{i}g\cos\theta(\partial^\mu\bar{c}^-)Z_\mu c^+ - \mathrm{i}e(\partial^\mu\bar{c}^+)W_\mu^-c^\gamma\\
&-\mathrm{i}g\cos\theta(\partial^\mu\bar{c}^+)W_\mu^-c^z + \mathrm{i}e(\partial^\mu\bar{c}^+)A_\mu c^- + \mathrm{i}g\cos\theta(\partial^\mu\bar{c}^+)Z_\mu c^-
\end{aligned}$$

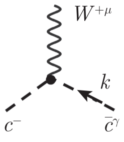
The relevant vertices are as follows:


$W^{+\mu}$
k
c^-
$\bar{c}^z$

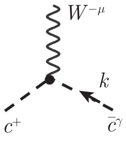
$\mathrm{i}g\cos\theta\ k^\mu$


$W^{-\mu}$
k
c^+
$\bar{c}^z$

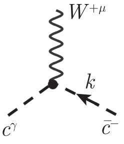
$-\mathrm{i}g\cos\theta\ k^\mu$


$W^{+\mu}$
k
c^-
$\bar{c}^\gamma$

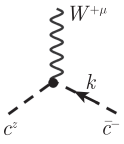
$\mathrm{i}ek^\mu$


$W^{-\mu}$
k
c^+
$\bar{c}^\gamma$

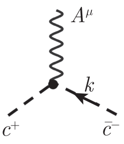
$-\mathrm{i}ek^\mu$


$W^{+\mu}$
k
c^γ
$\bar{c}^-$

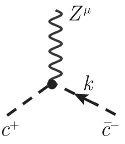
$-\mathrm{i}ek^\mu$


$W^{+\mu}$
k
c^z
$\bar{c}^-$

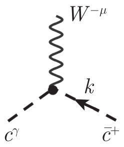
$-\mathrm{i}g\cos\theta\ k^\mu$



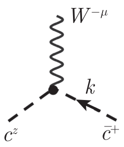
$\mathrm{i}ek^\mu$



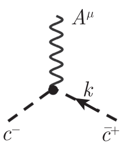
$\mathrm{i}g\cos\theta\ k^\mu$


$W^{-\mu}$
k
c^γ
$\bar{c}^+$

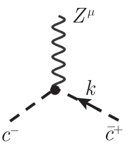
$\mathrm{i}ek^\mu$


$W^{-\mu}$
k
c^z
$\bar{c}^+$

$\mathrm{i}g\cos\theta\ k^\mu$



$-\mathrm{i}ek^\mu$

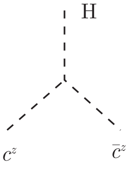


$-\mathrm{i}g\cos\theta\ k^\mu$

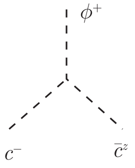
(ii) *The interaction vertex of ghost and scalar fields*:

$$-\frac{g}{2\cos\theta}m_Z\xi_Z\bar{c}^z c^z H+\frac{1}{2}gm_Z\xi_Z\bar{c}^z c^-\phi^+ +\frac{1}{2}gm_Z\xi_Z\bar{c}^z c^+\phi^-$$
$$-em_W\xi_W\bar{c}^- c^\gamma\phi^+ -\frac{g\cos 2\theta}{2\cos\theta}m_W\xi_W\bar{c}^- c^z\phi^+ -\frac{1}{2}gm_W\xi_W\bar{c}^- c^+ H$$
$$-\frac{\mathrm{i}}{2}gm_W\xi_W\bar{c}^- c^+\chi-em_W\xi_W\bar{c}^+ c^\gamma\phi^- -\frac{g\cos 2\theta}{2\cos\theta}m_W\xi_W\bar{c}^+ c^z\phi^-$$
$$-\frac{1}{2}gm_W\xi_W\bar{c}^+ c^- H+\frac{\mathrm{i}}{2}gm_W\xi_W\bar{c}^+ c^-\chi$$

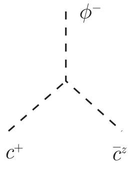
The relevant vertices are as follows:



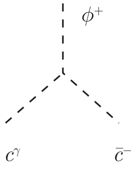
$-\mathrm{i}\dfrac{g}{2\cos\theta}m_Z\xi_Z$



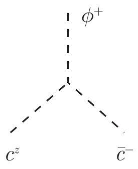
$\dfrac{\mathrm{i}}{2}gm_Z\xi_Z$



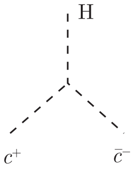
$\dfrac{\mathrm{i}}{2}gm_Z\xi_Z$


ϕ^+
c^γ
$\bar{c}^-$

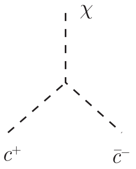
$-\mathrm{i}em_W\xi_W$



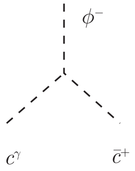
$-\mathrm{i}\dfrac{g\cos 2\theta}{2\cos\theta}m_W\xi_W$



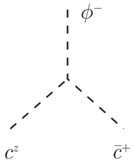
$-\dfrac{\mathrm{i}}{2}gm_W\xi_W$



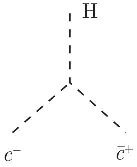
$\dfrac{1}{2}gm_W\xi_W$


ϕ^-
c^γ
$\bar{c}^+$

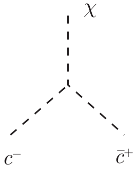
$-\mathrm{i}em_W\xi_W$



$-\mathrm{i}\dfrac{g\cos 2\theta}{2\cos\theta}m_W\xi_W$

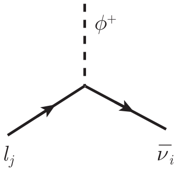


$-\dfrac{\mathrm{i}}{2}gm_W\xi_W$

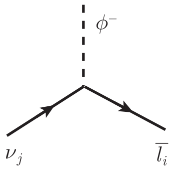


$-\dfrac{1}{2}gm_W\xi_W$

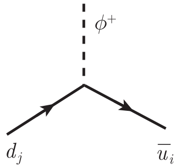
(iii) *The interaction vertex of fermion and scalar fields*: The coupling of $\phi^\pm$ and χ with fermion fields appear in the R_ξ gauge, which leads to new vertices. According to Eq. (8.8.34), the new added vertices are as follows:


ϕ^+
l_j
$\overline{\nu}_i$

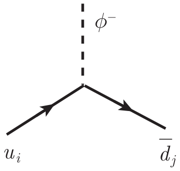
$$-\frac{\mathrm{i}m_{li}}{\sqrt{2}\upsilon}(1+\gamma_5)\delta_{ij}$$


ϕ^-
ν_j
$\overline{l}_i$

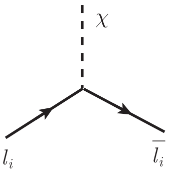
$$-\frac{\mathrm{i}m_{li}}{\sqrt{2}\upsilon}(1-\gamma_5)\delta_{ij}$$


ϕ^+
d_j
$\overline{u}_i$

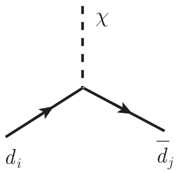
$$\frac{\mathrm{i}}{\sqrt{2}\upsilon}V_{ij}[(m_{ui}-m_{dj})-\gamma_5(m_{ui}+m_{dj})]$$


ϕ^-
u_i
$\overline{d}_j$

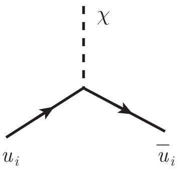
$$\frac{\mathrm{i}}{\sqrt{2}\upsilon}V_{ij}^*[(m_{ui}-m_{dj})+\gamma_5(m_{ui}+m_{dj})]$$


χ
l_i
$\overline{l}_i$

$$\frac{m_{li}}{\upsilon}\gamma_5$$


χ
d_i
$\overline{d}_j$

$$\frac{m_{di}}{\upsilon}\gamma_5$$


χ
u_i
$\overline{u}_i$

$$-\frac{m_{ui}}{\upsilon}\gamma_5$$

8.9. The Discovery of Higgs Boson

On July 4th, 2012, the European organization for nuclear research (CERN) announced the discovery of a new particle at the large hadron collider (LHC), which is very probably the Higgs particle. Its production and decay characteristics are consistent with that of Higgs boson of the standard model. The mass of the new particle measured by ATLAS collaboration working on LHC is 126.0 ± 0.4 (stat.) $\pm$ 0.4 (syst.) GeV, and the signal significance of the new particle over the background reaches 5.9 standard deviations [15], while the mass measured by CMS collaboration is 125.3 ± 0.4 (stat.) $\pm$ 0.5 (syst.) GeV, and the significance of the observation is 5.0 standard deviation [16].

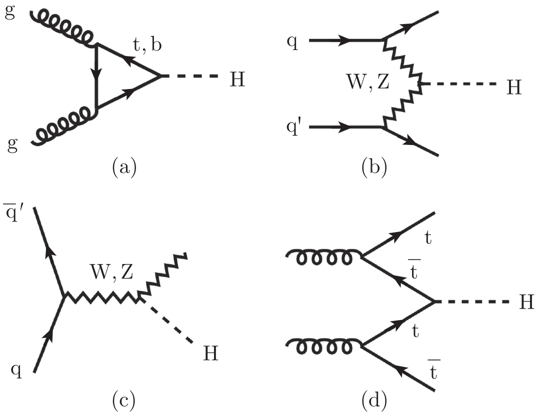


Fig. 8.3. Feynman diagrams for Higgs production in proton–proton collision in the standard model.

On March 14th, 2013, CERN released another news issue, announcing that analysis of more experimental data showed that the newly discovered particle is the Higgs boson.

LHC is a proton–proton collider. In the case of proton–proton collision, the processes of producing Higgs boson in the standard model are: (1) two-gluon fusion, gg → H, which is shown in Fig. 8.3(a); (2) vector boson fusion, $\mathrm{qq}' \to \mathrm{qq}'H$, as shown in Fig. 8.3(b); (3) Higgs radiation, $\mathrm{qq}' \to \mathrm{WH}, \mathrm{ZH}$, as shown in Fig. 8.3(c); (4) generation associated with $t\bar{t}$ production $\mathrm{gg} \to \mathrm{t\bar{t}H}$, which is shown in Fig. 8.3(d), and $\mathrm{q\bar{q}} \to \mathrm{t\bar{t}H}$. All these processes have been studied thoroughly in theory with higher-order QCD and electroweak corrections being included. In the case of Higgs mass $m_{\mathrm{H}} = 125\,\mathrm{GeV}$, the total cross section of producing Higgs in LHC predicted in the standard model is (i) $\sigma = 17.5\,\mathrm{pb}$ when the center-of-mass energy of proton–proton collision being $\sqrt{S} = 7\,\mathrm{TeV}$; (ii) $\sigma = 22.3\,\mathrm{pb}$ as the center-of-mass energy being $\sqrt{S} = 8\,\mathrm{TeV}$ [17,18].

In the standard model, there are five most promising decay modes for Higgs particle to be detected at LHC: (i) $\mathrm{H} \to \gamma\gamma$, which is a loop induced decay mode, and there is no tree level contribution for this process in the standard model; (ii) $\mathrm{H} \to \mathrm{ZZ}^*$, and $\mathrm{ZZ}^* \to (l^+l^-)(l^+l^-)$, which in shorthand is $\mathrm{H} \to \mathrm{ZZ}^* \to 4l$; (iii) $\mathrm{H} \to \mathrm{WW}^*$, and continually $\mathrm{WW}^* \to (l\bar{\nu})(l^+\nu)$, in shorthand it is $\mathrm{H} \to \mathrm{WW}^* \to 2l2\nu$; (iv) $\mathrm{H} \to \tau^-\tau^+$, where at least one lepton τ decays into pure leptonic final states; (v) $\mathrm{H} \to \mathrm{b\bar{b}}$, with b quark fragmenting into hadronic jet. Higgs decays in a very short time

Table 8.1. The branching ratios of various main decay modes of Higgs calculated in the standard model with $m_H = 125\,\text{GeV}$ [17].

$\text{H} \to \text{b}\bar{\text{b}}$	$\text{H} \to \tau^-\tau^+$	$\text{H} \to \text{c}\bar{\text{c}}$	$\text{H} \to \gamma\gamma$	$\text{H} \to \text{WW}$	$\text{H} \to \text{ZZ}$
57.8%	6.37%	2.68%	0.23%	21.6%	2.67%

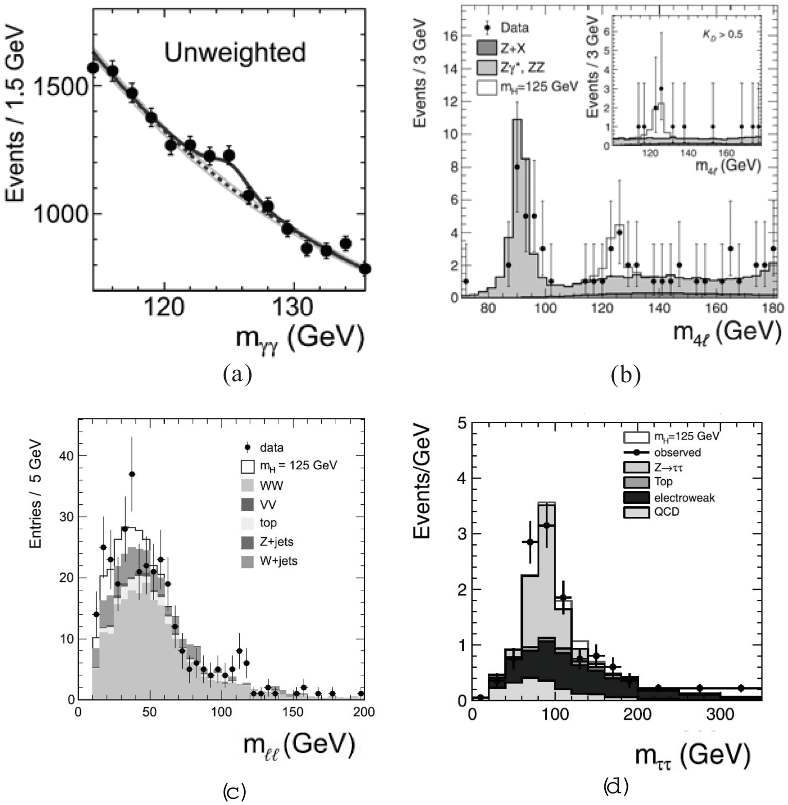


Fig. 8.4. The invariant mass distribution of Higgs measured by CMS collaboration [16]. (a) The invariant mass distribution of two photons $m_{\gamma\gamma}$ due to $\text{H} \to \gamma\gamma$ process; (b) the invariant mass distribution of leptons in $\text{H} \to \text{ZZ}^* \to 4l$ process; (c) the invariant mass distribution of two leptons m_{ll} in $\text{H} \to \text{WW}^* \to 2l2\nu$ process, where the solid curve is the expected distribution if a Higgs boson with mass $m_\text{H} = 125\,\text{GeV}$ exists; (d) the invariant mass distribution of τ leptons in $\text{H} \to \tau^+\tau^-$ process. The vertical axes in the figures are for event numbers detected in experiment in the relevant decay modes.

after its being generated. Its mean lifetime is very short. Theoretical calculation shows that, the total width of Higgs is $\Gamma_{\rm H} = 4.03 \times 10^{-3}\,{\rm GeV}$ [17] when its mass is taken to be $m_{\rm H} = 125\,{\rm GeV}$. The relevant lifetime is then $\tau_{\rm H} = \hbar/\Gamma_{\rm H} = 1.63 \times 10^{-22}\,{\rm s}$. The branching ratios of various decay modes of Higgs particle calculated in the standard model with $m_{\rm H} = 125\,{\rm GeV}$ are listed in Table 8.1.

The invariant mass distribution of Higgs boson measured by CMS collaboration is shown in Fig. 8.4 [16], which shows that there is apparent peak structure at $m_{\rm H} = 125\,{\rm GeV}$. The mass distribution of leptons m_{ll} in Fig. 8.4(c) deviates slightly from $m_{\rm H}$, that is because there are neutrinos in the final state in the process $\rm H \to WW^* \to 2\mathit{l}2\nu$, and neutrinos can carry away energy and momentum. But the distribution of m_{ll} is consistent with that expected from Higgs decays with $m_{\rm H} = 125\,{\rm GeV}$.

Figure 8.5 shows the ratio of Higgs production cross section times the branching ratio of the relevant Higgs decay mode over the theoretical prediction of the standard model [16]. The measured results in the figure show that the experimental data are consistent with the standard model within the error range.

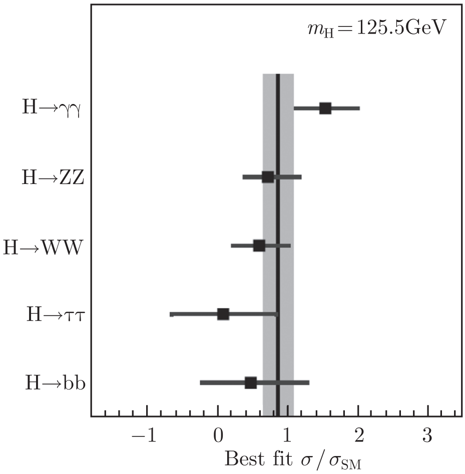


Fig. 8.5. The ratio $\sigma/\sigma_{\rm SM}$ for various decay modes [16]. $\sigma/\sigma_{\rm SM}$ denotes the ratio of Higgs production cross section times the branching ratio of the relevant Higgs decay mode over the theoretical prediction of the standard model. The gray band in the figure stands for the allowed range for the ratio $\sigma/\sigma_{\rm SM} = 0.87 \pm 0.23$. The horizontal error bars of the experimental data corresponding to the various decay modes represent the range of one standard deviation, where statistical and systematic errors are included.

Discovering Higgs is one of the important physical goals of LHC. After the discovery of the Higgs boson, the next step is to continue to study the properties of Higgs in experiment. This will help to understand the nature of microinteraction for particles and the source of mass for particles.

References

1. Yang C N, Mills R L. *Phys. Rev.*, 1954, 96: 191.
2. Glashow S L. *Nucl. Phys.*, 1961, 22: 579.
3. Weinberg S. *Phys. Rev. Lett.*, 1967, 19: 1264.
4. Salam A. *Elementary Particle Physics: Relativistic Groups and Analyticity*, Svartholm N, ed. *Proceedings of the Eighth Nobel Symposium.* Almqvist & Wiskell, 1968: 367.
5. Englert F, Brout R. *Phys. Rev. Lett.*, 1964, 13: 321.
6. Higgs P W. *Phys. Lett.*, 1964, 12: 132.
7. Higgs P W. *Phys. Rev. Lett.*, 1964, 13: 508.
8. Guralnik G S, Hagen C R, Kibble T W B. *Phys. Rev. Lett.*, 1964, 13: 585.
9. Higgs P W. *Phys. Rev.*, 1966, 145: 1156.
10. Kibble T W B. *Phys. Rev.*, 1967, 155: 1554.
11. Goldstone J. *Nuovo Cim.*, 1961, 19: 154.
12. Goldstone J, Salam A, Weinberg S. *Phys. Rev.*, 1962, 127: 965.
13. Faddeev L D, Popov V N. *Phys. Lett.*, 1967, B25: 29.
14. Muta T. *Foundations of Quantum Chromodynamics: An Introduction to Perturbative Methods in Gauge Theories.* World Scientific Publishing Co. Pte. Ltd, 1998.
15. Aad G, *et al.* (ATLAS Collab.). *Phys. Lett.*, 2012, B716: 1–29.
16. Chatrchyan S, *et al.* (CMS Collab.). *Phys. Lett.*, 2012, B716: 30–61.
17. Dittmaier S, *et al.* LHC Higgs Cross Section Working Group, Handbook of LHC Higgs Cross Section: 1. Inclusive observables, CERN-2011-002, 2011.
18. Dittmaier S, *et al.* LHC Higgs Cross Section Working Group, Handbook of LHC Higgs Cross Section: 2. Differential Distributions, CERN-2012-002, 2012.

Chapter 9

Theory of Strong Interaction: QCD

Quantum chromodynamics (QCD) is the theory of strong interaction between quarks. It is a theory based on symmetry of non-Abelian gauge group.

In 1967, a large electron linear accelerator was built at the Stanford Linear Accelerator Center (SLAC) in the United States of America. For the next few years, experimentalists have continually been doing electron–proton inelastic scattering experiment on this accelerator. It was finally found that with the increase of electron scattering angle, the cross section of the deep inelastic scattering was similar to that of electron scattered by a hard scattering center [1]. Inspired by these experimental results, J. Bjorken predicted the scaling phenomenon in 1968 when he used the method of current algebra [2]. Feynman then proposed the parton model [3], where proton was postulated to be a complex of partons, and the electron–proton deep inelastic scattering is the elastic scattering between electron and partons. Calculation showed that the scaling phenomenon can be obtained by the parton model. The parton model can be unified with quark model, and the partons are just the valence quarks and sea quarks. The scaling phenomenon indicates that partons inside nucleons can be regarded as quasi-free particles in the case of large momentum transfer, that is, the strength of the interaction between quarks decreases with the increase of the momentum transfer, which is called asymptotic freedom. So the interaction that binds quarks inside nucleons should possess the property of asymptotic freedom. In 1973, D. Gross, F. Wilczek and H.D. Politzer studied the characteristics of non-Abelian gauge field by using the renormalization group method. They found that SU(3) non-Abelian gauge field possesses the characteristic of asymptotic freedom [4,5]. This finding played a key role in the development of a theory describing strong interactions. It was gradually believed that strong interaction between quarks should be described by a non-Abelian gauge theory. This gauge field theory is based on SU(3) color symmetry which is called quantum chromodynamics (QCD). The color symmetry group is usually denoted by $\mathrm{SU}(3)_{\mathrm{c}}$. It has been shown theoretically and experimentally that quarks carry color as well as flavor quantum numbers. Experiments show that there are three kinds of colors for quarks in nature, denoted as R (red), G (green) and B (blue). It is assumed in QCD that color symmetry is a strict symmetry for strong interaction. Particles that transmit strong interactions are called gluons. There are eight kinds of gluons in the case of color number being three. Unlike photons in quantum electrodynamics, gluons carry color numbers or color charges, and they can interact with each other directly. This is an essential reason that results in the asymptotic freedom in non-Abelian gauge field. After QCD was set up, it was quickly applied to the deep inelastic scattering process to test whether QCD can result in Bjorken's scaling phenomenon. After taking higher-order corrections in QCD into account, it was found that the scaling behavior was actually slightly damaged, and the damage was in the form of logarithms [6–8]. Therefore, scaling behavior is only a rough result in QCD. The logarithmic violation of the scaling rule was finally verified by the subsequent muon–nucleon deep inelastic scattering experiment [9, 10], which is a strong support to QCD. In recent decades, QCD has been tested by large amount of experiments, especially in the perturbation region, theoretical predictions agree well with experiments.

9.1. The SU(3) Gauge Symmetry of Color

Let $\psi_i(x)$ stand for a quark field carrying color, where the subscript "i" denotes the kind of color, and $i = 1, \ldots, n$. Experiments show that there are three kinds of colors in strong interaction, i.e. $n = 3$. n colored quark fields constitute the basic representation of SU(n) group. Under the SU(n) transformation, we have

$$\psi_i'(x) = U_{ij}\psi_j(x) \tag{9.1.1}$$

where U_{ij} is the element of SU(n) group, which can be written as

$$U = \mathrm{e}^{-\mathrm{i}T^a\theta^a} \tag{9.1.2}$$

where T_a in Eq. (9.1.2) is the generator of SU(n) group. There are totally n^2-1 generators, i.e. $a = 1, \ldots, n^2-1$. The generators satisfy the following commutation relation

$$[T^a, T^b] = \mathrm{i} f^{abc} T^c \tag{9.1.3}$$

where f^{abc} is the structure constant of the group. In all the above and all of the following equations, unless otherwise specified, summation over the repeated indices is indicated.

The Lagrangian for free fermion field is

$$\mathcal{L} = \bar{\psi}_i (\mathrm{i}\gamma^\mu \partial_\mu - m)\psi_i \tag{9.1.4}$$

As long as the SU(n) group parameters θ^a in Eq. (9.1.2) do not depend on the space–time coordinate x, the Lagrangian of free fermion must be invariant under the SU(n) transformation given in Eq. (9.1.1). This is called the SU(n) global symmetry of free fermion. However, if the group transformation parameters are changed to be dependent on space–time coordinate, $\theta^a \to \theta^a(x)$, then Eq. (9.1.4) will not be invariant under the transformation given in Eq. (9.1.1), because $\partial U \neq 0$, extra terms appear along with the SU(n) transformation for Eq. (9.1.4). Such extra terms can only be subtracted by introducing gauge fields. Then we need to change the space–time derivative into covariant derivative $\partial_\mu \to D_\mu$, that is

$$D_\mu = \partial_\mu - \mathrm{i} g T^a A^a_\mu \tag{9.1.5}$$

where A^a_μ is the gauge field associated with the local gauge transformation for the color interaction, and g the coupling parameter between the gauge field and fermions. If the color indices are explicitly written, then Eq. (9.1.5) becomes

$$D^{ij}_\mu = \delta_{ij}\partial_\mu - \mathrm{i} g T^a_{ij} A^a_\mu \tag{9.1.6}$$

The generators in Eq. (9.1.6) are in their basic representation. After the gauge fields are introduced, the Lagrangian is changed to be

$$\begin{aligned}\mathcal{L} &= \bar{\psi}_i (\mathrm{i}\gamma^\mu D^{ij}_\mu - m\delta_{ij})\psi_j \\ &= \bar{\psi}(\mathrm{i}\gamma^\mu D_\mu - m)\psi \end{aligned} \tag{9.1.7}$$

The gauge transformation where the group transformation parameters θ^a depend on space–time coordinate is referred to as local gauge transformation. The gauge transformation of the gauge field A^a_μ can be derived according to the requirement of the local gauge invariance of Eq. (9.1.7). To keep the invariance of the Lagrangian in Eq. (9.1.7) under local gauge transformation, the only condition, under local gauge transformation, is that we must have

$$(D_\mu\psi)' = UD_\mu\psi \tag{9.1.8}$$

Substituting the definition of the covariant derivative D_μ into Eq. (9.1.8), we have

$$(\partial_\mu - \mathrm{i}gT^a A'^a_\mu)\psi' = U(\partial_\mu - \mathrm{i}gT^a A^a_\mu)\psi \tag{9.1.9}$$

Considering $\psi' = U\psi$, we can obtain from Eq. (9.1.9)

$$(U\partial_\mu + \partial_\mu U - \mathrm{i}gT^a A'^a_\mu U)\psi = U(\partial_\mu - \mathrm{i}gT^a A^a_\mu)\psi \tag{9.1.10}$$

Eliminating the same terms on both sides and simplifying it, we can get

$$T^a A'^a_\mu = UT^a A^a_\mu U^{-1} - \frac{\mathrm{i}}{g}(\partial_\mu U)U^{-1} \tag{9.1.11}$$

Under the local gauge transformation

$$\psi' = U\psi, \quad U = \mathrm{e}^{-\mathrm{i}T^a\theta^a(x)} \tag{9.1.12}$$

and the transformation of Eq. (9.1.11), it is obvious that $\bar{\psi}\gamma^\mu D_\mu\psi$ is invariant, therefore the Lagrangian in Eq. (9.1.7) is invariant.

The Lagrangian in Eq. (9.1.7) describes the interaction of fermions with gauge field and the kinetic energy of fermions. A complete Lagrangian describing the system of fermions and color gauge fields should also include the kinematic terms of the color gauge fields, which should be invariant under gauge transformation given in Eq. (9.1.11). Similar to the case of Abelian gauge field in Eq. (8.2.18), the field strength tensor of non-Abelian gauge filed is defined through the commutator of covariant derivative operators

$$[D_\mu, D_\nu] = -\mathrm{i}gT^a F^a_{\mu\nu} \tag{9.1.13}$$

Calculating the commutator in Eq. (9.1.13), we can obtain

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + gf^{abc}A^b_\mu A^c_\nu \tag{9.1.14}$$

Similar to the kinetic energy term of electromagnetic field, the Lagrangian of color gauge field is taken to be

$$-\frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} \tag{9.1.15}$$

It can be proven that Eq. (9.1.15) is invariant under gauge transformation. We can derive the gauge transformation of the covariant derivative operator at first. From Eq. (9.1.8), one can get

$$D'_\mu \psi' = U D_\mu \psi \tag{9.1.16}$$

Using $\psi' = U\psi$, then

$$D'_\mu U \psi = U D_\mu \psi \tag{9.1.17}$$

Equation (9.1.17) holds for any state of the fermion field ψ, so we should have the following equation for operators

$$D'_\mu U = U D_\mu \tag{9.1.18}$$

That is, under gauge transformation, the covariant derivative operator transforms in the following way:

$$D'_\mu = U D_\mu U^{-1} \tag{9.1.19}$$

According to Eqs. (9.1.19) and (9.1.13), we can derive the gauge transformation of the gauge field strength tensor $F^a_{\mu\nu}$

$$T^a F'^a_{\mu\nu} = U T^a F^a_{\mu\nu} U^{-1} \tag{9.1.20}$$

Using $\mathrm{Tr}[T^a T^b] = \frac{1}{2}\delta_{ab}$, the kinetic energy term for gauge field can be rewritten as

$$-\frac{1}{4} F^a_{\mu\nu} F^{a\mu\nu} = -\frac{1}{2}\mathrm{Tr}[T^a F^a_{\mu\nu} T^b F^{b\mu\nu}] \tag{9.1.21}$$

Under gauge transformation, Eq. (9.1.21) transforms as

$$\begin{aligned}
-\frac{1}{4} F'^a_{\mu\nu} F'^{a\mu\nu} &= -\frac{1}{2}\,\mathrm{Tr}[T^a F'^a_{\mu\nu} T^b F'^{b\mu\nu}] \\
&= -\frac{1}{2}\,\mathrm{Tr}[U T^a F^a_{\mu\nu} U^{-1} U T^b F^{b\mu\nu} U^{-1}] \\
&= -\frac{1}{2}\,\mathrm{Tr}[T^a F^a_{\mu\nu} T^b F^{b\mu\nu}] \\
&= -\frac{1}{4} F^a_{\mu\nu} F^{a\mu\nu}
\end{aligned} \tag{9.1.22}$$

Equation (9.1.22) indicates that the kinetic energy term of the gauge field is indeed invariant under gauge transformation.

The Lagrangian that is invariant under non-Abelian gauge transformation can be summarized as

$$\mathcal{L} = -\frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} + \bar{\psi}(\mathrm{i}\gamma^\mu D_\mu - m)\psi \tag{9.1.23}$$

The above Lagrangian is obtained according to local gauge invariance. However, one must keep in mind that, with only the requirement of gauge invariance, the possibility cannot be excluded that the higher-order terms composed of gauge field strength tensor $F^a_{\mu\nu}$ and quark field ψ may also exist. We can know later that such terms will violate the renormalizability of the theory. Therefore, from the requirement of renormalizability, such higher-order terms must be excluded. Another feature of non-Abelian theory different from Abelian theory is that there are self-interaction terms of gauge fields included in the Lagrangian in Eq. (9.1.23). These self-interaction terms are introduced by $gf^{abc}A^b_\mu A^c_\nu$ in $F^a_{\mu\nu}$, which does not exist in electromagnetic interaction. They are the source of asymptotic freedom of QCD.

Equation (9.1.23) is for the case that there is only one type of quark field in theory. If there are more than one type of quark fields, then the Lagrangian of QCD becomes

$$\mathcal{L} = -\frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} + \sum_{k=1}^{N_f}\bar{\psi}^k(\mathrm{i}\gamma^\mu D_\nu - m_k)\psi^k \tag{9.1.24}$$

The sum in Eq. (9.1.24) is over all types of quark fields, and the gauge field strength tensor and the covariant derivative are

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + gf^{abc}A^b_\mu A^c_\nu \tag{9.1.25}$$

$$D_\mu = \partial_\mu - \mathrm{i}gT^a A^a_\mu \tag{9.1.26}$$

respectively. For SU(3) gauge field $a = 1, 2, \ldots, 8$. There are eight kinds of color gauge fields.

9.2. The Quantization of Gauge Field and Fermion Field

9.2.1. *The path integral quantization*

For non-Abelian gauge field, the most appropriate method of quantization is path integral quantization [11]. We discuss path integral quantization in this section. To reduce the complexity of the problem, we consider a system composed of a pure scalar field, which is denoted as $\phi(x)$, and the mass of the scalar field is m. At first let us derive the expression of the n-point Green function in the path integral quantization. In the canonical quantization, the n-point Green function is the vacuum expectation value of the time-ordered n scalar field operators $\langle 0|\, T[\hat{\phi}(x_1)\hat{\phi}(x_2)\cdots\hat{\phi}(x_n)]\,|0\rangle$. Green function can be related to scattering matrix element or decaying amplitude of particles through reduction formula, therefore Green function is a very important function in quantum field theory.

Assume that the field operator $\hat{\phi}(x)$ is the operator in Heisenberg representation, and $\hat{H}$ is the Hamiltonian operator for the system. Then $\hat{\phi}(x)$ satisfies the Heisenberg equation

$$\frac{\partial\hat{\phi}(x)}{\partial t} = \mathrm{i}[\hat{H}, \hat{\phi}(x)] \tag{9.2.1}$$

The operator $\hat{\phi}(x)$ at any time t is related to that at time $t = 0$ by

$$\hat{\phi}(x) = \mathrm{e}^{\mathrm{i}\hat{H}t}\hat{\phi}(0, \vec{x})\mathrm{e}^{-\mathrm{i}\hat{H}t} \tag{9.2.2}$$

Define the eigenstate of the operator $\hat{\phi}(0, \vec{x})$ at time $t = 0$ is $|\phi(0, \vec{x})\rangle$, that is

$$\hat{\phi}(0, \vec{x})|\phi(0, \vec{x})\rangle = \phi(0, \vec{x})|\phi(0, \vec{x})\rangle \tag{9.2.3}$$

Then according to the above two equations, one can know that the eigenstate of $\hat{\phi}(x)$ is $\mathrm{e}^{\mathrm{i}\hat{H}t}|\phi(0, \vec{x})\rangle$, and the eigenvalue is $\phi(0, \vec{x})$.

Next let us analyze the transition matrix element between two eigenstates of scalar field operator. Let $|\phi_i, t_i\rangle$ and $|\phi_f, t_f\rangle$ be the initial and final states, respectively. Then the transition matrix element between the initial and final states of the scalar field $\phi(x)$ is

$$\langle \phi_f, t_f \mid \phi_i, t_i\rangle = \langle \phi_f|\, \mathrm{e}^{-\mathrm{i}\hat{H}(t_f - t_i)}\, |\phi_i\rangle \tag{9.2.4}$$

The field $\phi(x)$ distributes in a continuous infinite space–time volume. To treat the above matrix element, we restrict the distribution of the field in a limited space at first. Let V be the volume of the space. Divide this finite space into N small volume elements, with the volume of the small element being v, then $V = Nv$. After dealing with the transition matrix elements, we will take the limit $N \to \infty$, $v \to 0$, while keeping V unchanged. At last we shall take $V \to \infty$ to recover the infinite space. Take an arbitrary point $\vec{x}_j (j = 1, 2, \ldots, N)$ in each infinitesimal volume as a representative point, and the field in this infinitesimal volume is represented by the value at this representative point, then the continually distributed field in the whole space is divided into a set of discrete fields $\phi_j \equiv \phi(0, \vec{x}_j)$. The time interval between the initial and final states $t_f - t_i$ can also be divided into $(M+1)$ small time intervals. Let this small time interval to be Δ, and denote these time interval points as

$$t_0 = t_i, \quad t_1 = t_i + \Delta, \ldots, t_{M+1} = t_i + (M+1)\Delta = t_f \tag{9.2.5}$$

Denoting the field at time t_l $(l = 0, 1, 2, \ldots, M+1)$ and space coordinate $\vec{x}_j$ as ϕ_j^l, and considering the completeness of the field state

$$\int \mathrm{d}\phi_j^l \left|\phi_j^l, t_l\right\rangle \left\langle\phi_j^l, t_l\right| = 1 \tag{9.2.6}$$

we can write the transition matrix element as a product of a set of matrix elements in discrete space elements at first, then insert Eq. (9.2.6) into every matrix element in an increasing time-order from initial to final time,

$$\begin{aligned}\langle\phi_f, t_f|\, \phi_i, t_i\rangle &= \lim_{N\to\infty} \langle\phi_{f1}, \phi_{f2}, \ldots, \phi_{fN}, t_f|\, \phi_{i1}, \phi_{i2}, \ldots, \phi_{iN}, t_i\rangle \\ &= \lim_{N\to\infty} \prod_{j=1}^{N} \langle\phi_{fj}, t_f|\, \phi_{ij}, t_i\rangle \\ &= \lim_{\substack{N\to\infty \\ M\to\infty}} \int \prod_{j=1}^{N} \mathrm{d}\phi_j^M \cdots \mathrm{d}\phi_j^1 \,\langle\phi_{fj}, t_f|\, \phi_j^M, t_M\rangle \\ &\quad \cdot \left\langle\phi_j^M, t_M\right|\, \phi_j^{M-1}, t_{M-1}\big\rangle \cdots \left\langle\phi_j^1, t_1\right|\, \phi_{ij}, t_i\big\rangle\end{aligned} \tag{9.2.7}$$

When taking the number of time intervals to infinity $M \to \infty$ in the above equation, the time intervals shall tend to infinitesimal $\Delta \to 0$, so that to keep $M\Delta$ finite.

Considering the time evolution of the field state, one can expand each matrix element in Eq. (9.2.7) with respect to the time interval Δ

$$\begin{aligned}\left\langle\phi_j^{l+1}, t_{l+1}\right|\, \phi_j^l, t_l\big\rangle &= \left\langle\phi_j^{l+1}, t_l\right| \mathrm{e}^{-\mathrm{i}\hat{H}\Delta} \left|\phi_j^l, t_l\right\rangle \\ &= \left\langle\phi_j^{l+1}, t_l\right|\, \phi_j^l, t_l\big\rangle - \mathrm{i}\Delta \left\langle\phi_j^{l+1}, t_l\right| \hat{H} \left|\phi_j^l, t_l\right\rangle + O(\Delta^2)\end{aligned} \tag{9.2.8}$$

The Hamiltonian $\hat{H}$ can also be written in the form of a summation over all small volume elements

$$\hat{H} = \int \mathrm{d}^3x \hat{\mathcal{H}} = \sum_j \upsilon \hat{\mathcal{H}}_j \tag{9.2.9}$$

where $\hat{\mathcal{H}}_j$ is the Hamiltonian on the jth volume element. To treat the matrix element in Eq. (9.2.8) further, we expand the eigenstate of the field operator into momentum eigenstate. Let the momentum operator in the jth volume element is $\hat{p}_j$, whose eigenstate and eigenvalue are $|p_j^l\rangle$ and p_j^l, respectively. Then we have

$$\left\langle p_j^l \right| \left. \phi_j^l, t_l \right\rangle = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\mathrm{i} p_j^l \phi_j^l} \tag{9.2.10}$$

$$\left\langle p_j^l \right| \left. \phi_j^{l+1}, t_l \right\rangle = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\mathrm{i} p_j^l \phi_j^{l+1}} \tag{9.2.11}$$

Using the completeness of momentum eigenstates $\int \mathrm{d}p_j \, |p_j\rangle \langle p_j| = 1$, Eq. (9.2.8) becomes

$$\begin{aligned}
\left\langle \phi_j^{l+1}, t_{l+1} \right| \left. \phi_j^l, t_l \right\rangle &= \int \mathrm{d}p_j^l \left\langle \phi_j^{l+1}, t_l \right| \left. p_j^l \right\rangle \left\langle p_j^l \right| \left. \phi_j^l, t_l \right\rangle \\
&\quad - \mathrm{i}\Delta\upsilon \int \mathrm{d}p_j^l \mathrm{d}p_j'^l \left\langle \phi_j^{l+1}, t_l \right| \left. p_j'^l \right\rangle \left\langle p_j'^l \left| \hat{\mathcal{H}}_j \right| p_j^l \right\rangle \\
&\quad \times \left\langle p_j^l \right| \left. \phi_j^l, t_l \right\rangle + O(\Delta^2) \\
&= \int \mathrm{d}p_j^l \left\langle \phi_j^{l+1}, t_l \right| \left. p_j^l \right\rangle \left\langle p_j^l \right| \left. \phi_j^l, t_l \right\rangle \\
&\quad - \mathrm{i}\Delta\upsilon \int \mathrm{d}p_j^l \left\langle \phi_j^{l+1}, t_l \right| \left. p_j^l \right\rangle \mathcal{H}_j^l \left\langle p_j^l \right| \left. \phi_j^l, t_l \right\rangle + O(\Delta^2) \\
&= \int \frac{\mathrm{d}p_j^l}{2\pi} \mathrm{e}^{\mathrm{i} p_j^l \left(\phi_j^{l+1} - \phi_j^l\right)} \left(1 - \mathrm{i}\Delta\upsilon \mathcal{H}_j^l\right) + O(\Delta^2) \\
&= \int \frac{\mathrm{d}p_j^l}{2\pi} \mathrm{e}^{\mathrm{i}\left[p_j^l \left(\phi_j^{l+1} - \phi_j^l\right) - \Delta\upsilon \mathcal{H}_j^l\right]} + O(\Delta^2)
\end{aligned} \tag{9.2.12}$$

where $\mathcal{H}_j^l$ is the eigenvalue of the Hamiltonian operator in momentum eigenstate $\left|p_j^l\right\rangle$, and here the momentum eigenstates are also the eigenstates of the Hamiltonian operator. Therefore, we have $\left\langle p_j'^l \right| \hat{\mathcal{H}}_j \left| p_j^l \right\rangle = \mathcal{H}_j^l \delta(p_j^l - p_j'^l)$. Substituting Eq. (9.2.12) into Eq. (9.2.7), we can obtain

$$\begin{aligned}
\langle \phi_f, t_f | \, \phi_i, t_i \rangle &= \lim_{\substack{N\to\infty \\ M\to\infty}} \int \prod_{j=1}^{N} \left(\prod_{l=1}^{M} \mathrm{d}\phi_j^l \prod_{l'} \frac{\mathrm{d}p_j^{l'}}{2\pi} \right) \\
&\quad \times \exp\left\{ \mathrm{i} \sum_{l=0}^{M} \sum_{j=1}^{N} \left[p_j^l \left(\phi_j^{l+1} - \phi_j^l\right) - \Delta\upsilon \mathcal{H}_j^l \right] \right\}
\end{aligned} \tag{9.2.13}$$

Introducing momentum density π_j^l, which satisfies $p_j^l = \upsilon\pi_j^l$, then Eq. (9.2.13) becomes

$$\langle\phi_f, t_f|\ \phi_i, t_i\rangle = \lim_{\substack{N\to\infty \\ M\to\infty}} \int \prod_{j=1}^{N} \left(\prod_{l=1}^{M} \mathrm{d}\phi_j^l \prod_{l'} \frac{\upsilon \mathrm{d}\pi_j^{l'}}{2\pi}\right)$$
$$\times \exp\left\{\mathrm{i}\sum_{l=0}^{M} \Delta \sum_{j=1}^{N} \upsilon \left(\pi_j^l \frac{\phi_j^{l+1} - \phi_j^l}{\Delta} - \mathcal{H}_j^l\right)\right\} \tag{9.2.14}$$

Equation (9.2.14) can be written into a more compact form

$$\langle\phi_f, t_f|\ \phi_i, t_i\rangle$$
$$= \int [\mathrm{d}\phi] \left[\frac{\upsilon \mathrm{d}\pi}{2\pi}\right] \exp\left\{\mathrm{i}\int_{t_i}^{t_f} \mathrm{d}t \int \mathrm{d}^3x \left[\pi(x)\dot{\phi}(x) - \mathcal{H}(x)\right]\right\} \tag{9.2.15}$$

For the case of scalar field, substituting the following Hamiltonian

$$\mathcal{H} = \frac{1}{2}\pi^2 + \frac{1}{2}(\nabla\varphi)^2 + \frac{1}{2}m^2\phi^2 + V(\phi) \tag{9.2.16}$$

and integrating over the momentum π, we can obtain

$$\langle\phi_f, t_f|\ \phi_j, t_i\rangle = C\int [\mathrm{d}\phi] \exp\left(\mathrm{i}\int_{t_i}^{t_f} \mathrm{d}t \int \mathrm{d}^3x \mathcal{L}\right) \tag{9.2.17}$$

where $\mathcal{L}$ is the canonical Lagrangian for scalar field

$$\mathcal{L} = \frac{1}{2}\left(\partial^\mu\phi\partial_\mu\phi - m^2\phi^2\right) - V(\phi) \tag{9.2.18}$$

and C is the integral constant.

Equation (9.2.17) is the transition matrix element between two scalar field eigenstates. Next let us consider the matrix element $\langle\phi_f, t_f|\, \hat{\phi}(x)\, |\phi_i, t_i\rangle$, where a field operator $\hat{\phi}(x)$ is inserted between two field eigenstates. If the field operator $\hat{\phi}(x)$ is at time $t = t_l$, then according to the above analysis, we can obtain

$$\langle\phi_f, t_f|\, \hat{\phi}(x)\, |\phi_i, t_i\rangle = \lim_{\substack{N\to\infty \\ M\to\infty}} \int \left(\prod_{jl} \mathrm{d}\phi_j^l\right) \left\langle\phi_{fj}, t_f \mid \phi_j^M, t_M\right\rangle$$
$$\cdots \left\langle\phi_j^{l+1}, t_{l+1}\right| \hat{\phi}(x) \left|\phi_j^l, t_l\right\rangle$$
$$\cdots \left\langle\phi_j^1, t_1 \mid \phi_{ij}, t_i\right\rangle \tag{9.2.19}$$

Since $\left|\phi_j^l, t_l\right\rangle$ is the eigenstate of the field operator $\hat{\phi}(x)$, $\hat{\phi}(x)\left|\phi_j^l, t_l\right\rangle = \phi_j^l\left|\phi_j^l, t_l\right\rangle$, Eq. (9.2.19) can be finally written as

$$\langle\phi_f, t_f|\,\hat{\phi}(x)\,|\phi_i, t_i\rangle = C\int[\mathrm{d}\phi]\,\phi(x)\exp\left(\mathrm{i}\int_{t_i}^{t_f}\mathrm{d}t\int\mathrm{d}^3x\mathcal{L}\right) \tag{9.2.20}$$

Similarly one can derive the matrix element in the case that the product of two field operators $\hat{\phi}(x_1)$ and $\hat{\phi}(x_2)$ are inserted between two field eigenstates. For the case $t_1 = t_k > t_2 = t_l$, we can obtain

$$\begin{aligned}\langle\phi_f, t_f|\,\hat{\phi}(x_1)\hat{\phi}(x_2)\,|\phi_i, t_i\rangle = \lim_{\substack{N\to\infty\\ M\to\infty}}&\int\left(\prod_{jl}\mathrm{d}\phi_j^l\right)\left\langle\phi_{fj}, t_f \mid \phi_j^M, t_M\right\rangle\\ &\cdots\left\langle\phi_j^{k+1}, t_{k+1}\right|\hat{\phi}(x_1)\left|\phi_j^k, t_k\right\rangle\\ &\cdots\left\langle\phi_j^{l+1}, t_{l+1}\right|\hat{\phi}(x_2)\left|\phi_j^l, t_l\right\rangle\\ &\cdots\left\langle\phi_j^1, t_1 \mid \phi_{ij}, t_i\right\rangle\end{aligned} \tag{9.2.21}$$

Considering $\left|\phi_j^k, t_k\right\rangle$ and $\left|\phi_j^l, t_l\right\rangle$ are all eigenstates of the field operator, we can finally obtain

$$\begin{gathered}\langle\phi_f, t_f|\,\hat{\phi}(x_1)\hat{\phi}(x_2)\,|\phi_i, t_i\rangle = C\int[\mathrm{d}\phi]\,\phi(x_1)\phi(x_2)\exp\left(\mathrm{i}\int_{t_i}^{t_f}\mathrm{d}t\int\mathrm{d}^3x\mathcal{L}\right)\\ t_f > t_1 > t_2 > t_i\end{gathered} \tag{9.2.22}$$

In the same way, for the case of $t_2 > t_1$, the right-hand side of Eq. (9.2.22) stands for the matrix element $\langle\phi_f, t_f|\,\hat{\phi}(x_2)\hat{\phi}(x_1)|\phi_i, t_i\rangle$. Therefore, for arbitrary t_1 and t_2, we can obtain the following result

$$\begin{aligned}&\langle\phi_f, t_f|\,T\left[\hat{\phi}(x_1)\hat{\phi}(x_2)\right]|\phi_i, t_i\rangle\\ &\quad= C\int[\mathrm{d}\phi]\,\phi(x_1)\phi(x_2)\exp\left(\mathrm{i}\int_{t_i}^{t_f}\mathrm{d}t\int\mathrm{d}^3x\mathcal{L}\right)\end{aligned} \tag{9.2.23}$$

where the symbol T in Eq. (9.2.23) stands for time-ordered product

$$T\left[\hat{\phi}(x_1)\hat{\phi}(x_2)\right] = \theta\left(t_1 - t_2\right)\hat{\phi}(x_1)\hat{\phi}(x_2) + \theta\left(t_2 - t_1\right)\hat{\phi}(x_2)\hat{\phi}(x_1) \tag{9.2.24}$$

and $\theta(t)$ is the Heaviside step function. In general, for the matrix element of the time-ordered product of any more field operators, we have

$$\begin{aligned}&\langle\phi_f,t_f|\,T[\hat{\phi}(x_1)\cdots\hat{\phi}(x_n)]\,|\phi_i,t_i\rangle\\&\quad=C\int[\mathrm{d}\phi]\,\phi(x_1)\cdots\phi(x_n)\exp\left(\mathrm{i}\int_{t_i}^{t_f}\mathrm{d}t\int\mathrm{d}^3x\mathcal{L}\right)\end{aligned}\tag{9.2.25}$$

The time-ordered product of n field operators can be directly obtained by generalizing the time-ordered product of two field operators, which contains $n!$ terms.

In the following we shall consider the contribution of the ground state of the field to Eqs. (9.2.17) and (9.2.25). According to Eqs. (9.2.2) and (9.2.3), the eigenstate of the field operator $\hat{\phi}(x)$ is $\mathrm{e}^{\mathrm{i}\hat{H}t}\,|\phi(0,\vec{x})\rangle$, which is denoted as $|\phi(x)\rangle=\mathrm{e}^{\mathrm{i}\hat{H}t}\,|\phi(0,\vec{x})\rangle$. We can expand the eigenstate of the field operator $\hat{\phi}(x)$ on the eigenstates of the Hamiltonian, i.e. on the energetic eigenstates $|E_m\rangle$:

$$|\phi\,(x)\rangle=\sum_m\mathrm{e}^{\mathrm{i}E_mt}\,|E_m\rangle\,\langle E_m|\,\phi\,(0,\vec{x})\rangle\tag{9.2.26}$$

For the energetic eigenstate in Eq. (9.2.26), we have $\hat{H}\,|E_m\rangle=E_m\,|E_m\rangle$. As $m=0$, $|E_0\rangle$ is relevant to vacuum state $|0\rangle$ with $E_0=0$. Using Eq. (9.2.26), the matrix element on the left-hand side of Eq. (9.2.25) can be written as

$$\begin{aligned}&\langle\phi_f,t_f|T[\hat{\phi}(x_1)\cdots\hat{\phi}(x_n)]\,|\phi_i,t_i\rangle\\&\quad=\sum_{m,m'}\mathrm{e}^{\mathrm{i}E_mt_i-iE_{m'}t_f}\,\langle\phi_f\mid E_{m'}\rangle\,\langle E_m\mid\phi_i\rangle\,\langle E_{m'}|\,T[\hat{\phi}(x_1)\cdots\hat{\phi}(x_n)]\,|E_m\rangle\end{aligned}\tag{9.2.27}$$

Taking $t_i=\mathrm{i}T$, $t_f=-\mathrm{i}T$ in Eq. (9.2.27), and taking the limit $T\to\infty$, then on the right-hand side, except the ground state $E_m=E_{m'}=0$, the contributions of all the other terms tend to zero. Therefore, in the limit of $T\to\infty$, only the contribution of the vacuum state survives,

$$\begin{aligned}&\langle\phi_f,t_f|\,T[\hat{\phi}(x_1)\cdots\hat{\phi}(x_n)]\,|\phi_i,t_i\rangle\\&\quad\xrightarrow{T\to\infty}\langle\phi_f\mid 0\rangle\,\langle 0\mid\phi_i\rangle\,\langle 0|\,T[\hat{\phi}(x_1)\cdots\hat{\phi}(x_n)]\,|0\rangle\end{aligned}\tag{9.2.28}$$

where $t_i=\mathrm{i}T$, $t_f=-\mathrm{i}T$.

With the same treatment to Eq. (9.2.17), we can obtain

$$\langle \phi_f, t_f \mid \phi_i, t_i \rangle \xrightarrow{T\to\infty} \langle \phi_f \mid 0 \rangle \langle 0 \mid \phi_i \rangle \tag{9.2.29}$$

According to the above two equations and the results of Eqs. (9.2.17) and (9.2.25), we obtain

$$\begin{aligned}&\langle 0| T[\hat{\phi}(x_1)\cdots\hat{\phi}(x_n)] |0\rangle\\ &= \lim_{T\to\infty} \frac{\int [\mathrm{d}\phi]\, \phi(x_1)\cdots\phi(x_n) \exp\left(\mathrm{i}\int_{\mathrm{i}T}^{-\mathrm{i}T} \mathrm{d}t\mathrm{d}^3x\mathcal{L}\right)}{\int [\mathrm{d}\phi] \exp\left(\mathrm{i}\int_{\mathrm{i}T}^{-\mathrm{i}T} \mathrm{d}t\mathrm{d}^3x\mathcal{L}\right)}\end{aligned} \tag{9.2.30}$$

The time-integration in Eq. (9.2.30) is performed on the imaginary axis. Since there is no singular point for time variable in the integration, therefore, according to Cauchy's theorem, one can rotate 90° anticlockwise for the routine of time-integration back to real axis (see Fig. 9.1), then one obtain

$$\langle 0| T[\hat{\phi}(x_1)\cdots\hat{\phi}(x_n)] |0\rangle = \frac{\int [\mathrm{d}\phi]\, \phi(x_1)\cdots\phi(x_n) \exp(\mathrm{i}S)}{\int [\mathrm{d}\phi] \exp(\mathrm{i}S)} \tag{9.2.31}$$

where $S = \int \mathrm{d}^4x\mathcal{L}$, is the action of the system. This is the path integral for Green function.

Schwinger has reformed Eq. (9.2.31) into a more compact form by introducing the external source term [12]. Within this method, an external

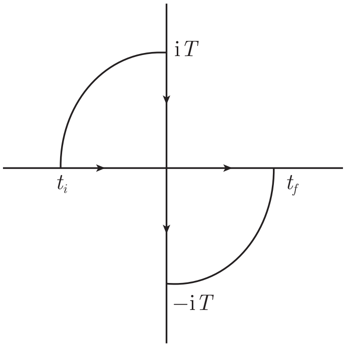


Fig. 9.1.

source function $J(x)$ has been artificially introduced, and an external source term $\phi(x)J(x)$ has been added in the functional integral in the denominator of Eq. (9.2.31), which is defined as generating functional

$$Z[J] = \int [\mathrm{d}\phi] \exp\left(\mathrm{i} \int \mathrm{d}^4 x\, (\mathcal{L} + \phi J)\right) \tag{9.2.32}$$

Generally $Z[J]$ is a function of the external source $J(x)$. Define the functional derivative of $Z[J]$ to J as

$$\frac{\delta Z[J]}{\delta J(y)} = \lim_{\varepsilon \to 0} \frac{Z[J(x) + \varepsilon\delta(x-y)] - Z[J(x)]}{\varepsilon} \tag{9.2.33}$$

By definition of functional derivative, we find that

$$\frac{\delta^n Z[J]}{\delta J(x_1)\cdots\delta J(x_n)} = \mathrm{i}^n \int [\mathrm{d}\phi]\, \phi(x_1)\cdots\phi(x_n) \exp\left\{\mathrm{i} \int \mathrm{d}^4 x\, (\mathcal{L} + \phi J)\right\} \tag{9.2.34}$$

and then, we can get

$$\langle 0|\, T[\hat{\phi}(x_1)\cdots\hat{\phi}(x_n)]\, |0\rangle = \frac{(-\mathrm{i})^n}{Z[0]} \left.\frac{\delta^n Z[J]}{\delta J(x_1)\cdots\delta J(x_n)}\right|_{J=0} \tag{9.2.35}$$

Equation (9.2.35) indicates that, all Green functions can be obtained by taking the functional derivative of $Z[J]$. It is in this sense that $Z[J]$ is called generating functional.

9.2.2. *Reduction formula*

Green function can be related to transition matrix element that describes particle scattering process. Therefore, the scattering cross section and decay width of particles can be calculated by calculating Green function at first. In this section, we derive the relationship between Green function and the scattering transition matrix element. The formula describing this relationship is called the reduction formula [13].

As a simple example, let's consider the elastic scattering of two scalar particles. Let the momenta of the initial state particles be p_1 and p_2, the momenta of the final state particles be k_1 and k_2, and all the masses of these particles be m. The S-matrix element of the scattering process is defined as

$$\langle k_1 k_2|\, S\, |p_1 p_2\rangle = \langle k_1 k_2\, \mathrm{out} \mid p_1 p_2\, \mathrm{in}\rangle \tag{9.2.36}$$

where "in" and "out" denote initial and final states before and after the interaction occurs during the scattering process, which are the eigenstates of the Hamiltonian of the system. Mathematically, the time of the initial state is taken as the infinite past and the time of the final state as the infinite future. Both the initial and final state particles can be regarded as free particles and can be represented by the creation operators acting on the vacuum state

$$|p_1p_2\,\mathrm{in}\rangle = \hat{a}^+_{\mathrm{in}}(p_1)\hat{a}^+_{\mathrm{in}}(p_2)\,|0\rangle \tag{9.2.37}$$

$$|k_1k_2\,\mathrm{out}\rangle = \hat{a}^+_{\mathrm{out}}(k_1)\hat{a}^+_{\mathrm{out}}(k_2)\,|0\rangle \tag{9.2.38}$$

Here $\hat{a}_{\mathrm{in}}$ and $\hat{a}_{\mathrm{out}}$ are all the coefficients in the Fourier transformation of the scalar fields for the initial and final states. Since the scalar field is an operator, these coefficients are also operators. They can be written uniformly by dropping the subscripts

$$\hat{\phi}(x) = \int \frac{\mathrm{d}^3p}{(2\pi)^3 2p_0}\left[\hat{a}(p)\mathrm{e}^{-\mathrm{i}p\cdot x} + \hat{a}^+(p)\mathrm{e}^{\mathrm{i}p\cdot x}\right] \tag{9.2.39}$$

Since both the initial and final state fields are free fields, the creation and annihilation operators satisfy the commutation relations

$$\begin{aligned}&\left[\hat{a}(p), \hat{a}^+(p')\right] = (2\pi)^3 2p_0\delta^3(p'-p)\\ &\left[\hat{a}(p), \hat{a}(p')\right] = 0\end{aligned} \tag{9.2.40}$$

According to Eq. (9.2.39), it is easy to prove

$$\hat{a}(p) = \mathrm{i}\int \mathrm{d}^3x\mathrm{e}^{\mathrm{i}p\cdot x}\overleftrightarrow{\partial}_0\hat{\phi}(x) \tag{9.2.41}$$

where the definition of the operator $\overleftrightarrow{\partial}_0$ in Eq. (9.2.41) is

$$f\overleftrightarrow{\partial}_0 g = f\partial_0 g - (\partial_0 f)\,g \tag{9.2.42}$$

The fields in the initial and final states $\hat{\phi}_{\mathrm{in}}(x)$ and $\hat{\phi}_{\mathrm{out}}(x)$ are defined as

$$\hat{\phi}(x) \xrightarrow[x_0\to\pm\infty]{} \hat{\phi}_{\substack{\mathrm{out}\\ \mathrm{in}}}(x) \tag{9.2.43}$$

With Eqs. (9.2.37), (9.2.38), (9.2.41) and (9.2.43) in hand, we can then derive the reduction formula

$$\begin{aligned}\langle k_1k_2\,\mathrm{out}\mid p_1p_2\,\mathrm{in}\rangle &= \langle k_1|\,\hat{a}_{\mathrm{out}}(k_2)\,|p_1p_2\,\mathrm{in}\rangle\\ &= \mathrm{i}\lim_{x_0\to+\infty}\int \mathrm{d}^3x\mathrm{e}^{\mathrm{i}k_2\cdot x}\overleftrightarrow{\partial}_0\,\langle k_1|\,\hat{\phi}(x)\,|p_1p_2\,\mathrm{in}\rangle\\ &= \langle k_1|\,\hat{a}_{\mathrm{in}}(k_2)\,|p_1p_2\,\mathrm{in}\rangle\\ &\quad+\mathrm{i}\int \mathrm{d}^4x\partial_0[\mathrm{e}^{\mathrm{i}k_2\cdot x}\overleftrightarrow{\partial}_0\,\langle k_1|\,\hat{\phi}(x)\,|p_1p_2\,\mathrm{in}\rangle]\end{aligned}\tag{9.2.44}$$

In the last step of Eq. (9.2.44), the following result has been used

$$\mathrm{i}\lim_{x_0\to-\infty}\int \mathrm{d}^3x\mathrm{e}^{\mathrm{i}k_2\cdot x}\overleftrightarrow{\partial}_0\hat{\phi}(x) = \hat{a}_{\mathrm{in}}(k_2)\tag{9.2.45}$$

Taking notice of $\partial_0(f\overleftrightarrow{\partial}_0 g) = f\overleftrightarrow{\partial}_0^2 g$ and $\partial_0^2\mathrm{e}^{\mathrm{i}k_2\cdot x} = \left(\nabla^2 - m^2\right)\mathrm{e}^{\mathrm{i}k_2\cdot x}$, one can further obtain

$$\begin{aligned}\langle k_1k_2\,\mathrm{out}\mid p_1p_2\,\mathrm{in}\rangle &= \langle k_1|\,\hat{a}_{\mathrm{in}}(k_2)\,|p_1p_2\,\mathrm{in}\rangle\\ &\quad+\mathrm{i}\int \mathrm{d}^4x\mathrm{e}^{\mathrm{i}k_2\cdot x}\left(\Box^2+m^2\right)\langle k_1|\,\hat{\phi}(x)\,|p_1p_2\,\mathrm{in}\rangle\end{aligned}\tag{9.2.46}$$

The operator $\Box$ in Eq. (9.2.46) is the d'Alembert operator, $\Box = \partial^\mu\partial_\mu = \partial^2/\partial t^2 - \nabla^2$. Repeat the above steps and one can finally get

$$\begin{aligned}&\langle k_1k_2\,\mathrm{out}\mid p_1p_2\,\mathrm{in}\rangle\\ &= \langle k_1k_2\,\mathrm{in}\mid p_1p_2\,\mathrm{in}\rangle + \mathrm{i}^4\int \mathrm{d}^4x_1\mathrm{d}^4x_2\mathrm{d}^4y_1\mathrm{d}^4y_2\mathrm{e}^{\mathrm{i}(k_1\cdot x_1+k_2\cdot x_2-p_1\cdot y_1-p_2\cdot y_2)}\\ &\quad\times\left(\Box_{x_1}+m^2\right)\left(\Box_{x_2}+m^2\right)\left(\Box_{y_1}+m^2\right)\left(\Box_{y_2}+m^2\right)\\ &\quad\times\langle 0|\,T[\hat{\phi}(x_1)\hat{\phi}(x_2)\hat{\phi}(y_1)\hat{\phi}(y_2)]\,|0\rangle\end{aligned}\tag{9.2.47}$$

Equation (9.2.47) is called reduction formula. The second term in Eq. (9.2.47) is the result of $\langle k_1k_2|\,S-1\,|p_1p_2\rangle$, which is relevant to the transition matrix element in the particle scattering process

$$\begin{aligned}&(2\pi)^4\mathrm{i}\delta^4\left(k_1+k_2-p_1-p_2\right)\langle k_1k_2|\,T\,|p_1p_2\rangle\\&\quad=\mathrm{i}^4\int \mathrm{d}^4x_1\mathrm{d}^4x_2\mathrm{d}^4y_1\mathrm{d}^4y_2\mathrm{e}^{\mathrm{i}(k_1\cdot x_1+k_2\cdot x_2-p_1\cdot y_1-p_2\cdot y_2)}\\&\qquad\times\left(\Box_{x_1}+m^2\right)\left(\Box_{x_2}+m^2\right)\left(\Box_{y_1}+m^2\right)\left(\Box_{y_2}+m^2\right)\\&\qquad\times\langle 0|\,T[\hat{\phi}(x_1)\hat{\phi}(x_2)\hat{\phi}(y_1)\hat{\phi}(y_2)]\,|0\rangle\end{aligned}\tag{9.2.48}$$

The same δ function as that in the left can appear on the right-hand side of the above equation after a few steps of mathematical manipulation, so we can compute the transition matrix element $\langle k_1k_2|\,T\,|p_1p_2\rangle$ with the above equation. With the transition matrix element in hand, we can then calculate the scattering cross section of any physical processes.

9.2.3. *The quantization of gauge field*

In the path integral quantization method, Green functions can be calculated from the generating functional, then from which one can calculate all possible physical quantities. Therefore, to discuss the quantization of the field, we only need to start with the discussion of the generating functional. In this section, we discuss the quantization of the canonical field from the generating functional.

According to the results of Eq. (9.2.31), for the case of gauge field, the corresponding generating functional is

$$Z\,[J]=\int[\mathrm{d}A]\exp\left\{\mathrm{i}\int \mathrm{d}^4x\left(\mathcal{L}+A_\mu^aJ^{a\mu}\right)\right\}\tag{9.2.49}$$

where $[\mathrm{d}A]$ stands for $\prod_{\mu,a}\left[\mathrm{d}A_\mu^a\right]$, and $\mathcal{L}$ is the Lagrangian for the gauge field

$$\mathcal{L}=-\frac{1}{4}F_{\mu\nu}^aF^{a\mu\nu}\tag{9.2.50}$$

In Eq. (9.2.49), $\mathcal{L}$ is gauge invariant, $[\mathrm{d}A]$ is also gauge invariant, but the external source term $A_\mu^aJ^{a\mu}$ is gauge-dependent. Therefore, the generation functional $Z\,[J]$ is not gauge invariant. The gauge-dependence of $Z\,[J]$ is associated with the gauge-dependence of Green functions. At first, let us prove the gauge invariance of $[\mathrm{d}A]$.

We consider the infinitesimal gauge transformation, i.e. the case that the transformation parameter θ^a in Eq. (9.1.11) and Eq. (9.1.12) is an infinitesimal parameter. Here for brevity, we have omitted the apparent dependence of θ^a on x, but actually it is still a function depending on space–time coordinates. According to Eqs. (9.1.11) and (9.1.12), take θ^a as a small parameter and retain to the first order of θ^a, the transformation of the gauge field under the infinitesimal gauge transformation can be obtained

$$A'^a_\mu = A^a_\mu + f^{abc}\theta^b A^c_\mu - \frac{1}{g}\partial_\mu \theta^a \tag{9.2.51}$$

So we can derive the transformation of the functional integral measure $[\mathrm{d}A]$ in $Z[J]$ under the infinitesimal gauge transformation

$$\begin{aligned}[\mathrm{d}A'] &= [\mathrm{d}A]\det\left(\frac{\partial A'^a_\mu}{\partial A^b_\nu}\right)\\ &= [\mathrm{d}A]\det\left(\delta^{ab} - f^{abc}\theta^c\right)\\ &= [\mathrm{d}A]\left(1 + O(\theta^2)\right)\end{aligned} \tag{9.2.52}$$

In the last step of the calculation in Eq. (9.2.52), the expansion about determinant is used

$$\det(1+L) = 1 + \mathrm{Tr}\, L + \cdots + \det L \tag{9.2.53}$$

Equation (9.2.52) shows that $[\mathrm{d}A]$ is gauge invariant if dropping terms of θ higher than the first power.

If remove the external source term, then the generating functional is gauge invariant. Let us use $Z[0]$ to discuss the problem of the quantization of the gauge field. It can be known from Eq. (9.2.49)

$$Z[0] = \int [\mathrm{d}A]\exp(\mathrm{i}S) \tag{9.2.54}$$

where $S = \int \mathrm{d}^4x\mathcal{L}$, is the action of the system, and the action is gauge-independent. In the functional integral of Eq. (9.2.54), the value of the gauge field A^a_μ varies in the whole space, that is, the integral of the gauge field is in the whole space of A^a_μ. But, because of the gauge symmetry, for a portion value of A^a_μ, the action S remains the same. Assume that the gauge field $A'^a_\mu = A^{(\theta)a}_\mu$ for a certain gauge parameter θ, here we explicitly write the dependence of the field on the gauge parameter. If we take the gauge transformation $U(\theta)$ for the gauge field from a particular value of A^a_μ, and let θ run over the whole range, then all the possible resulted $A^{(\theta)a}_\mu$ make up a subset. For all $A^{(\theta)a}_\mu$ in this subset, because of the gauge invariance of the action, the integrand in Eq. (9.2.54) remains as a constant. But this subset space is infinite, so the integral over this subspace is also infinite. Therefore, in Eq. (9.2.54), the contribution of $A_\mu^{(\theta)a}$ in this subset needs only to be taken once for the integral over $[\mathrm{d}A]$. The infinite factor of the subset space should be separated. To achieve this goal, we need to give the gauge field A_μ^a some restrictions. In principle, the restrictions can be arbitrarily selected

$$G^\mu A_\mu^a = B^a, \tag{9.2.55}$$

In Eq. (9.2.55), G^μ and B are the conditions selected to impose restriction on A_μ^a, which is called the gauge fixing condition. In general, we select the gauge condition such that there is a unique solution A_μ^a, and then the SU(3) group parameter θ^a corresponding to A_μ^a is determined. For example, when $G^\mu = \partial^\mu$, $B^a = 0$ are taken, such gauge is called Lorentz gauge, while the condition of $G^\mu = (0, \nabla)$, $B^a = 0$ is called Coulomb gauge.

To consider the restriction of the gauge condition given in Eq. (9.2.55) in the functional integral in the generating functional $Z[0]$, we define a functional $\Delta_G[A]$ in the following way:

$$\Delta_G[A] \int [\mathrm{d}g]\, \delta^n (G^\mu A_\mu^{(\theta)a} - B^a) = 1 \tag{9.2.56}$$

where the integration in Eq. (9.2.56) is performed in the whole space of the gauge group by running over all the group elements g, n is the dimension of the group, and $a = 1, 2, \ldots, n$. $[\mathrm{d}g]$ is the integral measure for the integration over the group elements, which can be selected in an expression over group parameter θ^a

$$[\mathrm{d}g] = \prod_a [\mathrm{d}\theta^a] \tag{9.2.57}$$

For a specific representation of the gauge group, one can always pick the appropriate group element integral measure $[\mathrm{d}g]$ to keep $[\mathrm{d}g]$ invariant under gauge transformation. Here we assume that the group integral measure $[\mathrm{d}g]$ in Eq. (9.2.56) has been taken as a gauge invariant form. It can be proved that the functional $\Delta_G[A]$ introduced by formula (9.2.56) is also gauge invariant. According to Eq. (9.2.56), we have

$$1/\Delta_G[A^{(\theta)}] = \int \prod_a [\mathrm{d}\theta'^a]\, \delta^n (G^\mu A_\mu^{(\theta\theta')a} - B^a) \tag{9.2.58}$$

Since the integral measure is invariant under gauge transformation, the integral measure in Eq. (9.2.58) can be transformed as $[\mathrm{d}\theta'^a] \to \left[\mathrm{d}\,(\theta\theta')^a\right]$, therefore Eq. (9.2.58) can be changed to be

$$\begin{aligned} 1/\Delta_G[A^{(\theta)}] &= \int \prod_a \left[\mathrm{d}\,(\theta\theta')^a\right] \delta^n(G^\mu A_\mu^{(\theta\theta')a} - B^a) \\ &= 1/\Delta_G\,[A] \end{aligned} \tag{9.2.59}$$

This indicates the functional $\Delta_G\,[A]$ is gauge invariant.

We now insert the left-hand side of Eq. (9.2.56) into the functional integral of Eq. (9.2.54). Since it is totally equal to 1, it does not change the value of the functional

$$Z\,[0] = \int [\mathrm{d}A] \prod_a [\mathrm{d}\theta^a]\, \delta^n(G^\mu A^{(\theta)a} - B^a)\Delta_G\,[A]\,\mathrm{e}^{\mathrm{i}S} \tag{9.2.60}$$

Changing the integral order with respect to group parameters and gauge field in Eq. (9.2.60), we have

$$Z\,[0] = \int \prod_a [\mathrm{d}\theta^a] \int [\mathrm{d}A]\, \delta^n(G^\mu A_\mu^{(\theta)a} - B^a)\Delta_G\,[A]\,\mathrm{e}^{\mathrm{i}S} \tag{9.2.61}$$

Equation (9.2.61) as a whole is gauge invariant, and the integral measure of the group parameter $\prod_a [\mathrm{d}\theta^a]$ is also gauge invariant, so the first integral in Eq. (9.2.61), i.e. the functional integral with respect to the gauge field $\int [\mathrm{d}A]$, must also be gauge invariant. So we can make gauge transformation for the integral for the gauge field, and select specific transformation parameter to make $A_\mu^{(\theta)a} \to A_\mu^a$. Under such a transformation, Eq. (9.2.61) does not change except that the variable of the δ function changed

$$Z\,[0] = \int \prod_a [\mathrm{d}\theta^a] \int [\mathrm{d}A]\, \delta^n\left(G^\mu A_\mu^a - B^a\right)\Delta_G\,[A]\,\mathrm{e}^{\mathrm{i}S} \tag{9.2.62}$$

At this point, the integrand of $Z\,[0]$ has no dependence on group parameters, so we can take the integral with respect to group parameters

$$\int \prod_a [\mathrm{d}\theta^a] \tag{9.2.63}$$

as an independent factor and separate it out. The factor is actually an infinite factor. Now we redefine the part after separating out the infinite factor as the new generating functional of the gauge field

$$Z\,[0] = \int [\mathrm{d}A]\,\Delta_G\,[A]\,\delta^n\left(G^\mu A^a_\mu - B^a\right)\mathrm{e}^{\mathrm{i}S} \tag{9.2.64}$$

Now let us figure out the concrete form of the functional $\Delta_G\,[A]$. Do the functional integral in Eq. (9.2.56):

$$\begin{aligned}1 &= \Delta_G\,[A]\int\prod_a[\mathrm{d}\theta^a]\,\delta^n\left(G^\mu A^{(\theta)a}_\mu - B^a\right)\\ &= \Delta_G[A]/\det M_G\end{aligned} \tag{9.2.65}$$

where the matrix M_G is obtained with manipulating the composite generalized function

$$(M_G(x,y))^{ab} = \frac{\delta(G^\mu A^{(\theta)a}_\mu(x))}{\delta\theta^b(y)} \tag{9.2.66}$$

Considering the result of $\Delta_G\,[A]$, and adding the external source term in Eq. (9.2.64), we obtain the generating functional $Z\,[J]$

$$\begin{aligned}Z\,[J] = \int [\mathrm{d}A]\det M_G\prod_x\delta^n\left(G^\mu A^a_\mu(x) - B^a(x)\right)\\ \times\exp\left\{\mathrm{i}\int\mathrm{d}^4x\left(\mathcal{L} + A^a_\mu J^\mu\right)\right\}\end{aligned} \tag{9.2.67}$$

Since $B^a(x)$ can be arbitrarily selected as a constraint, we can average the generating functional $Z\,[J]$ against $B^a(x)$ and take an appropriate weighting factor, which in principle can be arbitrarily selected. Here we take

$$\exp\left\{-\,(\mathrm{i}/2\alpha)\int\mathrm{d}^4x\,(B^a(x))^2\right\} \tag{9.2.68}$$

where α is an arbitrary constant, then multiplying this factor to the right-hand side of Eq. (9.2.67), and performing a functional integral with respect to $B^a(x)$, we can get

$$Z\,[J] = \int[\mathrm{d}A]\det M_G\exp\left\{\mathrm{i}\int\mathrm{d}^4x\left(\mathcal{L} - \frac{1}{2\alpha}\left(G^\mu A^a_\mu\right)^2 + A^a_\mu J^{a\mu}\right)\right\} \tag{9.2.69}$$

Using the above trick, the constraint by the δ function is converted to an exponential factor, which is generally referred to as gauge fixing term, and α is referred to as gauge fixing parameter.

The emergence of det M_G is significantly meaningful for the quantization of gauge field. It appears as a new term in the Lagrangian of gauge field, which will be discussed in more detail later. In the following, as examples, we give some results of M_G under a few gauge conditions. The results can be calculated directly by Eq. (9.2.66), only with a factor $-1/g$ omitted.

(1) Coulomb gauge $G^\mu = (0, \nabla)$

$$(M_G(x,y))^{ab} = \left(\delta^{ab}\nabla^2 - gf^{abc}A^c \cdot \nabla\right)\delta^4(x-y) \tag{9.2.70}$$

(2) Lorentz gauge $G^\mu = \partial^\mu$

$$(M_G(x,y))^{ab} = \left(\delta^{ab}\Box - gf^{abc}\partial^\mu A^c_\mu\right)\delta^4(x-y) \tag{9.2.71}$$

Note that the differential operator in Eq. (9.2.71) should also act on the δ function at the end of the equation.

(3) Axial gauge $G^\mu = n^\mu$, and n^μ is a space-like 4-vector

$$(M_G(x,y))^{ab} = \left(\delta^{ab} n \cdot \partial - gf^{abc} n \cdot A^c\right)\delta^4(x-y) \tag{9.2.72}$$

From the above results, we can see that, for Abelian gauge field $f^{abc} = 0$, M_G is a constant which is independent on the gauge field, so no new interaction term is added. For non-Abelian gauge field, especially in the case of using Lorentz gauge, M_G depends on gauge field, and it will bring new interaction term involving gauge field. Below we will exponentiate M_G into a form, which can be easily viewed as part of the Lagrangian of the system. This would lead to the introduction of an artificial field, called Faddeev–Popov's ghost field [14].

9.2.4. *The quantization of fermion field and Grassmann algebra*

In canonical quantization, the fermion field operators are anti-commutative, and in the path integral quantization, it is also necessary to reflect the anti-commutativity of fermion field. In the path integral formula involving gauge field, the values of the fields are described by common complex numbers, and ordinary complex numbers are commutative to each other, which is consistent with the commutativity of the boson fields. But to describe the fermion field in the path integral formula, using common complex number cannot reflect the anti-commutativity of fermionic field. To describe fermionic field in the path integral formulas, a new mathematical tool, Grassmann algebra, needs to be introduced.

Suppose that there are a new set of numbers ψ_j $(j = 1, 2, \ldots, N)$, which are different from ordinary complex numbers. They are anti-commutating complex numbers

$$\{\psi_i, \psi_j\} = 0 \tag{9.2.73}$$

These N anti-commutating numbers span an algebraic space called Grassmann algebra, and these anti-commutating numbers are called Grassmann numbers, which are commutative with ordinary numbers. Grassmann algebra can be thought of as a linear space consisting of a series of linearly independent basis vectors, which are composed of products of Grassmann numbers:

$$1, \psi_i, \psi_i\psi_j, \ldots, \psi_1\psi_2\cdots\psi_N \tag{9.2.74}$$

The number of these basis vectors is finite, and the amount is related to how many Grassmann number there are. If there are N Grassmann number for a system, then the basis vectors consist of at most N Grassmann number products. There must be repeated numbers in the product of more than N Grassmann number products, so such products must be 0, because for any Grassmann number ψ_i, due to the anti-commutativity, there must be $\psi_i^2 = 0$.

The number of basis vectors for a Grassmann algebra is equal to the total number of possible non-zero product terms, and the number of the non-zero product terms are

$$1 + \mathrm{C}_N^1 + \mathrm{C}_N^2 + \cdots + \mathrm{C}_N^N = 2^N \tag{9.2.75}$$

In Eq. (9.2.75), C_N^i is a binomial coefficient, so Grassmann algebra is 2^N-dimensional algebraic space.

For boson field, we divide the whole space into a series of small volume element. The field ϕ_j on each representative point in a small element is commuting complex number. In the case of fermion field, we can treat the field ψ_j on the representative point in each small volume as Grassmann number. Such that the anti-commutativity of fermion field is correctly described. The number of these discrete volume elements tends to infinity $N \to \infty$, so that the fermion field $\psi(x)$ can be described in the path integral formula. To see clearly the meaning of the integral for anti-commuting numbers in the path integral formula, we first introduce the differential and integral operations for Grassmann Numbers.

The differential of a single Grassmann number is defined as follows:

$$\frac{\partial}{\partial\psi_j}\psi_i = \delta_{ij} \tag{9.2.76}$$

For the differential of the product of several Grassmann numbers, for example, differentiating $\psi_{i_1}\psi_{i_2}\cdots\psi_{i_k}$ with respect to ψ_{i_l}, if there is ψ_{i_l} included in the product, then moving ψ_{i_l} to the most left part using the anti-commutativity at first, and then doing the differentiation with respect to ψ_{i_l}, which is relevant to moving ψ_{i_l} to the far left and then remove it; If there is no ψ_{i_l} in the product, the result of the differentiation is zero, that is, we have the following equation:

$$\frac{\partial}{\partial\psi_j}\left(\psi_{i_1}\psi_{i_2}\cdots\psi_{i_k}\right) = \begin{cases} (-1)^{l-1}\psi_{i_1}\psi_{i_2}\cdots(\psi_{i_l})\cdots\psi_{i_k} & \text{if } i_l = j \\ 0 & \text{if there is no } \psi_j \end{cases} \tag{9.2.77}$$

where (ψ_{i_l}) stands for that ψ_{i_l} is no longer contained in the expression. We can define the differentiation from the right in a similar manner, that is, moving ψ_{i_l} to the far right, then differentiating it

$$\left(\psi_{i_1}\psi_{i_2}\cdots\psi_{i_k}\right)\frac{\overleftarrow{\partial}}{\partial\psi_j} = \begin{cases} (-1)^{k-l}\psi_{i_1}\psi_{i_2}\cdots(\psi_{i_l})\cdots\psi_{i_k} & \text{if } i_l = j \\ 0 & \text{if there is no } \psi_j \end{cases} \tag{9.2.78}$$

Using Eq. (9.2.77), it can be proved that the second derivative of an arbitrary function F with respect to Grassmann number satisfies the following relation

$$\frac{\partial^2 F}{\partial\psi_i\partial\psi_j} = -\frac{\partial^2 F}{\partial\psi_j\partial\psi_i} \tag{9.2.79}$$

The proof is as follows.

Suppose F is composed of monomials constructed by the product of several Grassmann numbers. For one of the monomial $\psi_{i_1}\cdots\psi_{i_k}$, we have

$$\frac{\partial}{\partial\psi_i}\frac{\partial}{\partial\psi_j}(\psi_{i_1}\cdots\psi_{i_k}) = (-1)^{m-1}(-1)^{l-1}\psi_{i_1} \cdots(\psi_{i_l})\cdots(\psi_{i_m})\cdots\psi_{i_k} \tag{9.2.80}$$

$$\frac{\partial}{\partial\psi_j}\frac{\partial}{\partial\psi_i}(\psi_{i_1}\cdots\psi_{i_k}) = (-1)^{l-1}(-1)^{m-2}\psi_{i_1} \cdots(\psi_{i_l})\cdots(\psi_{i_m})\cdots\psi_{i_k} \tag{9.2.81}$$

We have taken $i_l = i$ and $i_m = j$ in the above two equations. From these two equations, one can directly obtain the result of Eq. (9.2.79). As a special case of (9.2.79), we have

$$\frac{\partial^2 F}{\partial\psi_i^2} = 0 \tag{9.2.82}$$

The definition of the integral for Grassmann numbers is

$$\int \mathrm{d}\psi_i = 0, \quad \int \mathrm{d}\psi_i\psi_j = \delta_{ij} \tag{9.2.83}$$

where

$$\{\mathrm{d}\psi_i, \psi_j\} = 0, \quad \{\mathrm{d}\psi_i, \mathrm{d}\psi_j\} = 0 \tag{9.2.84}$$

It should be noted that there is no reciprocal relation between the integral rule defined above and the differentiation of Grassmann number. For Grassmann number, integration and differentiation are two independent operations.

With the rules of differentiation and integration, we can perform all kinds of differential and integral operations on anti-commuting numbers. Let us do an integral about anti-commuting numbers that is very useful in path integral quantization. Compute the following integral:

$$I = \int \mathrm{d}\psi_1\mathrm{d}\psi_2\cdots\mathrm{d}\psi_N\mathrm{d}\bar{\psi}_1\mathrm{d}\bar{\psi}_2\cdots\mathrm{d}\bar{\psi}_N \exp(\bar{\psi}_i A_{ij}\psi_j) \tag{9.2.85}$$

In Eq. (9.2.85) sum over repeated indices is understood. A_{ij}'s are complex number, ψ_i and $\bar{\psi}_i$ $(i = 1, 2, \ldots, N)$ are all Grassmann numbers. $\bar{\psi}_i$ can be either understood as the complex conjugate of ψ_i, or new anti-commuting numbers independent on ψ_i. Because of the anti-commutativity of Grassmann numbers, $\psi_i^2 = 0$, $\bar{\psi}_i^2 = 0$, it only needs to expand the integrand in Eq. (9.2.85) to Nth order, and the terms of order higher than N are all 0

$$\exp(\bar{\psi}_i A_{ij}\psi_j) = 1 + \bar{\psi}_i A_{ij}\psi_j + \cdots + \frac{1}{N!}(\bar{\psi}_i A_{ij}\psi_j)^N \tag{9.2.86}$$

Then according to the integral rules in Eq. (9.2.83), only the integral of the term of order N is not zero, i.e.

$$I = \int \prod_{i=1}^{N} \mathrm{d}\psi_i \prod_{j=1}^{N} \mathrm{d}\bar{\psi}_j \frac{1}{N!} (\bar{\psi}_i A_{ij} \psi_j)^N \tag{9.2.87}$$

According to the anti-commutativity of Grassmann numbers, the power-N term in Eq. (9.2.87) can be calculated further

$$\begin{aligned}(\bar{\psi}_i A_{ij} \psi_j)^N &= (-1)^{N(N-1)/2} \sum_{i_1 \cdots i_N, j_1 \cdots j_N} A_{i_1 j_1} \cdots A_{i_N j_N} \bar{\psi}_{i_1} \\ &\quad \cdots \bar{\psi}_{i_N} \psi_{j_1} \cdots \psi_{j_N} \\ &= (-1)^{N(N-1)/2} \sum_{i_1 \cdots i_N, j_1 \cdots j_N} \varepsilon_{i_1 \cdots i_N} \varepsilon_{j_1 \cdots j_N} A_{i_1 j_1} \cdots A_{i_N j_N} \\ &\quad \times \bar{\psi}_1 \cdots \bar{\psi}_N \psi_1 \cdots \psi_N \end{aligned} \tag{9.2.88}$$

where $\varepsilon_{i_1 \cdots i_N}$ is antisymmetric tensor. Using the properties of antisymmetric tensor, one can further simplify Eq. (9.2.88) to be

$$\begin{aligned}(\bar{\psi}_i A_{ij} \psi_j)^N = (-1)^{N(N-1)/2} N! \sum_{i_1 \cdots i_N} \varepsilon_{i_1 \cdots i_N} A_{1 i_1} \\ \cdots A_{N i_N} \bar{\psi}_1 \cdots \bar{\psi}_N \psi_1 \cdots \psi_N \end{aligned} \tag{9.2.89}$$

Since the determinant of the matrix $A = (A_{ij})$ is

$$\det A = \sum_{i_1 \cdots i_N} \varepsilon_{i_1 \cdots i_N} A_{1 i_1} \cdots A_{N i_N} \tag{9.2.90}$$

so

$$(\bar{\psi}_i A_{ij} \psi_j)^N = (-1)^{N(N-1)/2} N! \det A \bar{\psi}_1 \cdots \bar{\psi}_N \psi_1 \cdots \psi_N \tag{9.2.91}$$

Substituting Eq. (9.2.91) into (9.2.87), and integrating with respect to ψ_i and $\bar{\psi}_j$, we can obtain

$$I = (-1)^{N(N-1)/2} \det A \tag{9.2.92}$$

Going back to the continuum limit $N \to \infty$, from Eq. (9.2.92), we obtain the functional integral of anti-commuting numbers

$$\int [\mathrm{d}\psi] \, [\mathrm{d}\bar{\psi}] \exp\left\{ \int \mathrm{d}^4 x \mathrm{d}^4 y \bar{\psi}(x) A(x, y) \psi(y) \right\} = \det A \tag{9.2.93}$$

There may be a difference of negative sign in Eq. (9.2.93), but such a sign can be canceled when this result is used to Green function expressed by the functional integral.

With the expression of the functional integral for Grassmann numbers, Eq. (9.2.93), we can extend the generating functional in Eq. (9.2.69) to including fermions. For this purpose, we introduce the external sources η and $\bar{\eta}$ for fermion fields at first. The source functions are all anti-commuting numbers. The relevant external source terms are $\bar{\psi}\eta$ and $\bar{\eta}\psi$. After including fermions, the generating functional of QCD is

$$\begin{aligned} Z\left[J,\eta,\bar{\eta}\right] = \int [\mathrm{d}A]\,[\mathrm{d}\psi]\,\left[\mathrm{d}\bar{\psi}\right]\det M_G \\ \times \exp\left\{\mathrm{i}\int \mathrm{d}^4x\left(\mathcal{L}-\frac{1}{2\alpha}\left(\partial^\mu A_\mu^a\right)^2 + A_\mu^a J^{a\mu} + \bar{\psi}\eta + \bar{\eta}\psi\right)\right\} \end{aligned} \tag{9.2.94}$$

where the Lagrangian is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^a F^{a\mu\nu} + \bar{\psi}\left(\mathrm{i}\gamma^\mu D_\mu - m\right)\psi \tag{9.2.95}$$

The Green function of fermion fields can be obtained by taking the functional derivative with respect to η and $\bar{\eta}$, which is similar to the case of the boson fields. It should be noted that ψ, $\bar{\psi}$, η and $\bar{\eta}$ are all anti-commuting Grassmann numbers, and it is necessary to pay attention to their anti-commutativity in calculating the functional derivative

$$\left\{\psi(x),\bar{\psi}(y)\right\} = 0, \quad \{\psi(x),\eta(y)\} = 0, \quad \{\eta,\bar{\eta}\} = 0,\ldots \tag{9.2.96}$$

Therefore, when defining the Green function of the fermion field, we should pay attention to the negative sign brought by the anti-commutative relation. For example, the relationship between the two-point Green function for fermion fields and the derivative of the generating functional should be

$$\langle 0|\,T[\hat{\psi}_\alpha(x)\hat{\bar{\psi}}_\beta(y)]\,|0\rangle = \frac{(-\mathrm{i})^2}{Z\left[0,0,0\right]}\left.\frac{\delta^2 Z\left[J,\eta,\bar{\eta}\right]}{\delta\bar{\eta}_\alpha(x)\delta\left(-\eta_\beta(y)\right)}\right|_{J=\bar{\eta}=\eta=0} \tag{9.2.97}$$

9.3. The Effective Lagrangian of QCD and Perturbation Theory

9.3.1. *The effective Lagrangian of QCD*

In addition to the Lagrangian, gauge fixing and external source terms in the generating functional of QCD, Eq. (9.2.94), there is the determinant $\det M_G$ resulted from the quantization of the gauge field. The appearance of this term makes it not so simple to deal with the interaction of gauge field. From Eqs. (9.2.70)–(9.2.72), we see that $\det M_G$ depends on the gauge field in both Coulomb and Lorentz gauges, except that it is a constant in the axial gauge ($n \cdot A = 0$). Therefore, in the gauges that $\det M_G$ depending on the gauge field, one must deal with the determinant term $\det M_G$ when considering the interactions involving gauge fields. In Lorentz gauge $G_\mu = \partial_\mu$, the theory is apparently covariant. This is the advantage of Lorentz gauge. So Lorentz gauge is also called the covariant gauge. Next we shall discuss the effective Lagrangian of QCD in the covariant gauge and develop perturbation expansion method to obtain the Feynman rule.

From Eq. (9.2.93), we know that for any determinant $\det A$, we can always write it into a functional integral form of anti-commuting Grassmann numbers. Using this relation, we introduce two anti-commuting fictional fields at first, which are denoted as $\chi^a(x)$ and $\chi^b(y)$. These two fields obey Grassmann algebra, and they are anti-commutative fields as fermions, but later we shall see that their propagators are in the form of bosons. Because of such strange properties, they are called Faddeev–Popov's ghost field [14]. By introducing ghost fields, $\det M_G$ can be written as an exponential functional integral:

$$\det M_G = \int [\mathrm{d}\chi]\,[\mathrm{d}\bar{\chi}] \exp\left\{-\mathrm{i}\int \mathrm{d}^4x\mathrm{d}^4y\bar{\chi}^a(x)\,(M_G(x,y))^{ab}\,\chi^b(y)\right\} \tag{9.3.1}$$

Here $\chi^{a,b}(x)$ are understood to be complex Grassmann numbers, and $\chi^a(x)$ and $\bar{\chi}^a(x)$ understood as two independent fields. Certainly one can directly select two independent real fields χ_1^a and χ_2^a instead, and do the same thing to χ^b too. Substituting $M_G(x,y)^{ab} = \left(\delta^{ab}\Box - gf^{abc}\partial^\mu A_\mu^c\right)\delta^4(x-y)$ in the covariant gauge, i.e. Eq. (9.2.71), into (9.3.1), and performing the integration by parts, we get

$$\det M_G = \int [\mathrm{d}\chi]\,[\mathrm{d}\bar{\chi}] \exp\left\{\mathrm{i}\int \mathrm{d}^4x\,(\partial^\mu\bar{\chi}^a(x))\,D_\mu^{ab}\chi^b(x)\right\} \tag{9.3.2}$$

where D_μ^{ab} is the covariant derivative in adjoint representation. Since the generator of SU(3) group in adjoint representation is $(T^a)_{bc} = -\mathrm{i}f^{abc}$, so

$$D_\mu^{ab} = \delta^{ab}\partial_\mu - gf^{abc}A_\mu^c \tag{9.3.3}$$

Substituting Eq. (9.3.2) into Eq. (9.2.94), the generating functional of QCD, we obtain

$$\begin{aligned} Z\left[J,\xi,\bar{\xi},\eta,\bar{\eta}\right] = &\int [\mathrm{d}A]\,[\mathrm{d}\chi]\,[\mathrm{d}\bar{\chi}]\,[\mathrm{d}\psi]\,\left[\mathrm{d}\bar{\psi}\right] \\ &\times \exp\left\{\mathrm{i}\int \mathrm{d}^4x\left(\mathcal{L} + AJ + \bar{\chi}\xi + \bar{\xi}\chi + \bar{\psi}\eta + \bar{\eta}\psi\right)\right\} \end{aligned} \tag{9.3.4}$$

where the external sources ξ^a and $\bar{\xi}^a$ for ghost fields have been introduced, both of which are Grassmann numbers, and AJ, $\bar{\chi}\xi$, $\bar{\xi}\chi$ are abbreviations in the following:

$$AJ = A_\mu^a J^{a\mu}, \quad \bar{\chi}\xi = \bar{\chi}^a\xi^a, \quad \bar{\xi}\chi = \bar{\xi}^a\chi^a \tag{9.3.5}$$

$\mathcal{L}$ in Eq. (9.3.4) is the effective Lagrangian of QCD, which contains the effect of the determinant $\det M_G$, and $\mathcal{L}$ can be written in the following form:

$$\mathcal{L} = \mathcal{L}_{\mathrm{G}} + \mathcal{L}_{\mathrm{GF}} + \mathcal{L}_{\mathrm{FP}} + \mathcal{L}_{\mathrm{F}} \tag{9.3.6}$$

$$\mathcal{L}_{\mathrm{G}} = -\tfrac{1}{4}F_{\mu\nu}^a F^{a\mu\nu}, \quad F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + gf^{abc}A_\mu^b A_\nu^c \tag{9.3.7}$$

$$\mathcal{L}_{\mathrm{GF}} = -\tfrac{1}{2\alpha}\left(\partial^\mu A_\mu^a\right)^2 \tag{9.3.8}$$

$$\mathcal{L}_{\mathrm{FP}} = \left(\partial^\mu \bar{\chi}^a\right) D_\mu^{ab}\chi^b \tag{9.3.9}$$

$$\mathcal{L}_{\mathrm{F}} = \bar{\psi}^i\left(\mathrm{i}\gamma^\mu D_\mu^{ij} - m\delta^{ij}\right)\psi^j \tag{9.3.10}$$

where i, j are color indices, and $D_\mu^{ij} = \delta^{ij}\partial_\mu - \mathrm{i}gT_{ij}^a A_\mu^a$ is the covariant derivative in the basis representation of SU(3) group. Only the case of one type of fermion is given here, but it is easy to generalize it to the case of more types of fermions. In the case of more fermions, it only needs to change Eq. (9.3.10) to sum over all quark fields. $\mathcal{L}_{\mathrm{G}}$ is the term for pure gauge field, $\mathcal{L}_{\mathrm{GF}}$ the gauge fixing term, $\mathcal{L}_{\mathrm{FP}}$ the Faddeev–Popov's ghost field term, and $\mathcal{L}_{\mathrm{F}}$ is the term for fermion. Equation (9.3.6) is the effective Lagrangian of QCD in the covariant gauge.

9.3.2. *Perturbation expansion*

Now let us discuss the perturbation theory of QCD based on the Lagrangian given in Eqs. (9.3.6) and (9.3.7)–(9.3.10). We treat the coupling constant in the Lagrangian g as a small parameter, and expand the generating functional of QCD in Eq. (9.3.4) with respect to it. In order to more easily see the main steps of the perturbation expansion, as an example, we first introduce the perturbation expansion of the theory of ϕ^3 about scalar field. Although this is not a theory that describes the physical reality, but as an example, it is very succinct.

9.3.2.1. *The perturbative expansion of ϕ^3 theory*

The Lagrangian of scalar ϕ^3 theory is

$$\mathcal{L} = \frac{1}{2}\left(\partial^\mu\phi\partial_\mu\phi - m^2\phi^2\right) - V(\phi) \tag{9.3.11}$$

$$V(\phi) = -\frac{g}{3!}\phi^3 \tag{9.3.12}$$

where g is the cubic coupling constant for scalar field, and the factor 1/3! is introduced for conciseness in the following processing of the theory. We divide $\mathcal{L}$ into free and interaction parts, where the interaction part refers to the part that involves the coupling constant g

$$\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_\mathrm{I} \tag{9.3.13}$$

$$\mathcal{L}_0 = \frac{1}{2}\left(\partial^\mu\phi\partial_\mu\phi - m^2\phi^2\right) \tag{9.3.14}$$

$$\mathcal{L}_\mathrm{I} = -V(\phi) = \frac{g}{3!}\phi^3 \tag{9.3.15}$$

where $\mathcal{L}_0$ is the Lagrangian for free field, and $\mathcal{L}_\mathrm{I}$ for the interaction part. The generating functional is

$$\begin{aligned} Z[J] &= \int [\mathrm{d}\phi] \exp\left\{\mathrm{i}\int \mathrm{d}^4x\,(\mathcal{L}_0 + \mathcal{L}_\mathrm{I} + \phi J)\right\} \\ &= \int [\mathrm{d}\phi] \exp\left\{\mathrm{i}\int \mathrm{d}^4x \mathcal{L}_\mathrm{I}\right\} \exp\left\{\mathrm{i}\int \mathrm{d}^4x\,(\mathcal{L}_0 + \phi J)\right\} \end{aligned} \tag{9.3.16}$$

Doing Taylor expansion for the interaction part in Eq. (9.3.16)

$$\exp\left\{\mathrm{i}\int \mathrm{d}^4x\mathcal{L}_\mathrm{I}\right\} = \sum_n \frac{1}{n!}\left(\mathrm{i}\int \mathrm{d}^4x\mathcal{L}_\mathrm{I}\right)^n = \sum_n \frac{1}{n!}\left\{\mathrm{i}\int \mathrm{d}^4x\left(\frac{g}{3!}\phi^3\right)\right\}^n \tag{9.3.17}$$

Substituting the above expansion back into Eq. (9.3.16), one can obtain the generating functional

$$\begin{aligned}Z[J] &= \int[\mathrm{d}\phi]\sum_n \frac{1}{n!}\left\{\mathrm{i}\int \mathrm{d}^4x\left(\frac{g}{3!}\phi^3\right)\right\}^n \exp\left\{\mathrm{i}\int \mathrm{d}^4x\,(\mathcal{L}_0+\phi J)\right\}\\ &= \int[\mathrm{d}\phi]\sum_n \frac{1}{n!}\left\{\mathrm{i}\int \mathrm{d}^4x\frac{g}{3!}\left(-\mathrm{i}\frac{\delta}{\delta J(x)}\right)^3\right\}^n \exp\left\{\mathrm{i}\int \mathrm{d}^4x\,(\mathcal{L}_0+\phi J)\right\}\\ &= \int[\mathrm{d}\phi]\sum_n \frac{1}{n!}\left\{\mathrm{i}\int \mathrm{d}^4x\mathcal{L}_\mathrm{I}\left(-\mathrm{i}\frac{\delta}{\delta J(x)}\right)\right\}^n \exp\left\{\mathrm{i}\int \mathrm{d}^4x\,(\mathcal{L}_0+\phi J)\right\}\\ &= \sum_n \frac{1}{n!}\left\{\mathrm{i}\int \mathrm{d}^4x\mathcal{L}_\mathrm{I}\left(-\mathrm{i}\frac{\delta}{\delta J(x)}\right)\right\}^n \int[\mathrm{d}\phi]\exp\left\{\mathrm{i}\int \mathrm{d}^4x\,(\mathcal{L}_0+\phi J)\right\}\end{aligned} \tag{9.3.18}$$

Functional differentiation has been used in the second step of Eq. (9.3.18). If the generation functional for free field is defined as

$$Z_0[J] = \int[\mathrm{d}\phi]\exp\left\{\mathrm{i}\int \mathrm{d}^4x\,(\mathcal{L}_0+\phi J)\right\} \tag{9.3.19}$$

then Eq. (9.3.18) can be expressed as

$$Z[J] = \sum_n \frac{1}{n!}\left\{\mathrm{i}\int \mathrm{d}^4x\mathcal{L}_\mathrm{I}\left(-\mathrm{i}\frac{\delta}{\delta J(x)}\right)\right\}^n Z_0[J] \tag{9.3.20}$$

Equation (9.3.20) is the perturbation expansion of the generating functional, which is independent on the specific expression of $\mathcal{L}_\mathrm{I}$, and it is a general expression. The proof can be achieved from the generalization of Eq. (9.3.18).

Considering the concrete form of $\mathcal{L}_\mathrm{I}$ in ϕ^3 theory, we can write explicitly the first few terms of Eq. (9.3.20)

$$\begin{aligned}Z[J] = \Bigg\{1 - \frac{g}{3!}\int \mathrm{d}^4x\left(\frac{\delta}{\delta J(x)}\right)^3 + \frac{1}{2}\left(\frac{g}{3!}\right)^2\Bigg(\int \mathrm{d}^4x\\ \times\left(\frac{\delta}{\delta J(x)}\right)^3\Bigg)^2 + \cdots\Bigg\} Z_0[J]\end{aligned} \tag{9.3.21}$$

For the convenience of finding the functional differential with respect to $J(x)$, we perform the functional integral over ϕ in $Z_0[J]$ at first. $Z_0[J]$ can be expressed into

$$Z_0\left[J\right] = \int \left[\mathrm{d}\phi\right] \exp\left\{\frac{-\mathrm{i}}{2}\int \mathrm{d}^4x\mathrm{d}^4y\phi(x)K(x,y)\phi(y) + \mathrm{i}\int \mathrm{d}^4x\phi(x)J(x)\right\} \tag{9.3.22}$$

where $K(x,y)$ is

$$K(x,y) = \delta^4(x-y)\left(\Box_y + m^2\right), \quad \Box_y = \frac{\partial}{\partial y^\mu}\frac{\partial}{\partial y_\mu} \tag{9.3.23}$$

To finish the integration over $\phi(x)$ in Eq. (9.3.22), we write Eq. (9.3.22) in the form based on a discrete space

$$Z_0\left[J\right] = \lim_{N\to\infty}\int \mathrm{d}\phi_1 \cdots \mathrm{d}\phi_N \exp\left(\frac{-\mathrm{i}}{2}\sum_{i,j}\phi_i K_{ij}\phi_j + \mathrm{i}\sum_i \phi_i J_i\right) \tag{9.3.24}$$

where K_{ij} is the matrix element resulted from the values of $K(x,y)$ in the discrete space, and $K_{ij} \to K(x,y)$ in the limit of the continuous space. Making the variable substitution, $\phi'_i = \phi_i - \sum_j (K^{-1})_{ij}J_j$, we can obtain

$$Z_0\left[J\right] = \lim_{N\to\infty}\exp\left\{\mathrm{i}\frac{1}{2}\sum_{i,j}J_i(K^{-1})_{ij}J_j\right\} I_N \tag{9.3.25}$$

$$I_N = \int \mathrm{d}\phi'_1 \cdots \mathrm{d}\phi'_N \exp\left(\frac{-\mathrm{i}}{2}\sum_{ij}\phi'_i K_{ij}\phi'_j\right) \tag{9.3.26}$$

where I_N is a functional integral independent of the external source J, and the integral can be done by diagonalize the matrix (K_{ij}). The result of the integration is a Gaussian integral, which is a constant. Dropping this irrelevant constant, we get

$$Z_0\left[J\right] = \exp\left\{\frac{\mathrm{i}}{2}\int \mathrm{d}^4x\mathrm{d}^4yJ(x)\Delta(x,y)J(y)\right\} \tag{9.3.27}$$

where $\Delta(x,y)$ is the inverse of $K(x,y)$, i.e.

$$\int \mathrm{d}^4zK(x,z)\Delta(z,y) = \delta^4(x-y) \tag{9.3.28}$$

One can obtain through Eq. (9.3.28)

$$\Delta(x,y) = -\int \frac{\mathrm{d}^4k}{(2\pi)^4}\frac{\mathrm{e}^{-\mathrm{i}k\cdot(x-y)}}{k^2 - m^2 + \mathrm{i}\varepsilon} \tag{9.3.29}$$

Equation (9.3.29) is the free propagator of the scalar field, where ε is a positive infinitesimal, which is introduced for guaranteeing the convergence of the Gaussian integral in Eq. (9.3.22). Equation (9.3.29) is also the two-point Green function for free scalar field, i.e.

$$\langle 0|\, T[\hat{\phi}(x)\hat{\phi}(y)]\, |0\rangle = \frac{(-\mathrm{i})^2}{Z_0\,[0]}\, \frac{\delta^2 Z_0\,[J]}{\delta J(x)\delta J(y)}\bigg|_{J=0} = -\mathrm{i}\Delta(x,y) \qquad (9.3.30)$$

With Eq. (9.3.27), substituting it into (9.3.21), one can obtain the generating functional $Z\,[J]$ order by order. Here we give the result of $Z\,[J]$ expanded up to the first order of g

$$Z\,[J] = Z_0\,[J]\,\{1 - gZ_1\,[J] + \cdots\} \qquad (9.3.31)$$

where

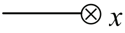
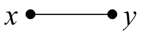
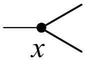
$$Z_1\,[J] = \frac{\mathrm{i}}{2!}\int \mathrm{d}^4x_1\Delta_{11}(\Delta_1 J_1) + \frac{\mathrm{i}^3}{3!}\int \mathrm{d}^4x_1(\Delta_1 J_1)(\Delta_1 J_2)(\Delta_1 J_3) \quad (9.3.32)$$

here Δ_{ij} and $(\Delta_i J_j)$ are defined as follows:

$$\Delta_{ij} = \Delta(x_i, x_j), \quad (\Delta_i J_j) = \int \mathrm{d}^4y_j\Delta(x_i, y_j)J(y_j) \qquad (9.3.33)$$

where i and j are integers.

Equation (9.3.32) can be expressed graphically. For this purpose, we make the following correspondence between graph and formulas:

$\int dxJ(x)$	$\leftrightarrow$		external source
$i\Delta(x,y)$	$\leftrightarrow$		propagator
$-\int d^4xg$	$\leftrightarrow$		vertex

Therefore, Eq. (9.3.32) can be expressed graphically as Fig. 9.2.

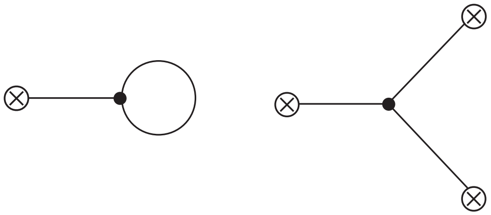


Fig. 9.2. The graphical expression of Eq. (9.3.32).

To determine the Green function, we have to determine $Z\left[0\right]$. Substituting $J=0$ into Eq. (9.3.31), we can obtain the result up to the first order of the coupling constant g

$$Z\left[0\right]=1+O(g)+\cdots \tag{9.3.34}$$

then we obtain

$$Z\left[J\right]/Z\left[0\right]=Z_0\left[J\right]\left\{1-gZ_1\left[J\right]+\cdots\right\} \tag{9.3.35}$$

After more tedious calculations, we can also compute $Z\left[J\right]/Z\left[0\right]$ up to the higher orders of the coupling constant, so as to use Eq. (9.2.35) to calculate the Green function to much higher orders of g. As mentioned above, we can use graphics to express the calculated results for Green function, and extract the rules that associate the numerical results with graphics, which are referred to as the Feynman rules, and the corresponding diagrams are called Feynman diagrams. Of course with Eq. (9.3.35), part of the Feynman diagrams corresponding to the calculated Green function are disconnected, which means that some Feynman diagrams are broken diagrams. However, the Feynman diagrams for the amplitude of physical scattering and decay processes are required to be connected. Meanwhile, it can be seen from the reduction formula that, the Feynman diagrams corresponding to physical amplitude are also what with external propagators being cut off. In mathematics, the generating functional corresponding to connected diagrams can be defined on the basis of $Z\left[J\right]$, which is called the generating functional of the connected Green function, denoted as $W\left[J\right]$. The definition is

$$W\left[J\right]=-\mathrm{i}\ln Z\left[J\right] \tag{9.3.36}$$

Finding the functional differentiation of $W\left[J\right]$, we can obtain the connected Green function

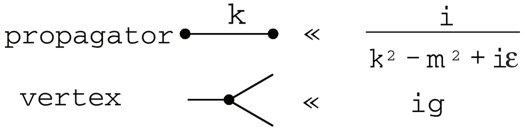
$$G_n^c(x_1,\ldots,x_n)=(-\mathrm{i})^{n-1}\left.\frac{\delta^n W\left[J\right]}{\delta J(x_1)\cdots\delta J(x_n)}\right|_{J=0} \tag{9.3.37}$$

The Feynman diagrams corresponding to Green function calculated by Eq. (9.3.37) are all connected, and the proof of this point is omitted.

The Feynman rules in the momentum space for the scalar ϕ^3 theory obtained by perturbative expansion are given below.

(a) The transition matrix element or scattering amplitude $\langle\cdots|\,T\,|\cdots\rangle$ corresponds to the connected, amputated Green function. When calculating the connected, amputated Green function of N external fields to order g^n, one needs to draw all topologically non-equivalent Feynman diagrams containing N external fields and n vertices.

(b) The following correspondence should be made



(c) Integrate for every closed loop diagram $\int \frac{\mathrm{d}^4k}{(2\pi)^4}$, here k is the free momentum in the loop line.

(d) Multiply a symmetric factor $1/S$, where $S = P\prod_{n=2,3}(n!)^{d_n}$ to each closed loop diagram, where d_n is the number of vertex pairs connected by n lines, and P all the possible permutation number of the vertices that keeps the Feynman diagram unchanged.

(e) Keep the four-momentum conserved in each vertices.

9.3.2.2. *The perturbative expansion of QCD and Feynman rule*

With the perturbation expansion method of scalar ϕ^3 theory as the basis, we can now discuss the perturbation expansion of QCD. We divide the Lagrangian of QCD into free field and interaction part

$$\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_\mathrm{I} \tag{9.3.38}$$

where the free field lagrangian $\mathcal{L}_0$ is what in Eq. (9.3.6) with the interaction between the gauge field and quark field being removed, that is, the part containing the coupling constant g being removed. The interaction Lagrangian $\mathcal{L}_\mathrm{I}$ refers to the interaction between gauge field and quark field in Eq. (9.3.6).

There are three parts in $\mathcal{L}_0$: the part for pure gauge field or gluon kinetic energy part, kinetic energy parts of Faddeev–Popov's ghost field and quark field

$$\mathcal{L}_0 = \mathcal{L}_0^\mathrm{G} + \mathcal{L}_0^\mathrm{FP} + \mathcal{L}_0^\mathrm{F} \tag{9.3.39}$$

$$\mathcal{L}_0^\mathrm{G} = -\frac{1}{4}\left(\partial_\mu A_\nu^a - \partial_\nu A_\mu^a\right)\left(\partial^\mu A^{a\nu} - \partial^\nu A^{a\mu}\right) - \frac{1}{2\alpha}\left(\partial^\mu A_\mu^a\right)^2 \tag{9.3.40}$$

$$\mathcal{L}_0^\mathrm{FP} = \left(\partial^\mu \bar{\chi}^a\right)\left(\partial_\mu \chi^a\right) \tag{9.3.41}$$

$$\mathcal{L}_0^\mathrm{F} = \bar{\psi}\left(\mathrm{i}\gamma^\mu\partial_\mu - m\right)\psi \tag{9.3.42}$$

It deserves to be noted that, according to Eq. (9.3.41), we can see that the Lagrangian for ghost is the same as the kinetic energy term of a massless charged (complex) boson field, thus ghost field has the characteristics of boson in behavior, but previously we have known that ghost field satisfies Grassmann algebra and is anti-commutative.

The interaction part of the Lagrangian $\mathcal{L}_{\mathrm{I}}$ contains all terms involving the coupling constant

$$\begin{aligned}\mathcal{L}_{\mathrm{I}} = &-\frac{g}{2} f^{abc}\left(\partial_\mu A_\nu^a - \partial_\nu A_\mu^a\right) A^{b\mu} A^{c\nu} - \frac{g^2}{4} f^{abe} f^{cde} A_\mu^a A_\nu^b A^{c\mu} A^{d\nu} \\ &- g f^{abc}\left(\partial^\mu \bar{\chi}^a\right) \chi^b A_\mu^c + g\bar{\psi} T^a \gamma_\mu \psi A^{a\mu} \end{aligned} \tag{9.3.43}$$

The perturbative expansion for the generating functional of the scalar ϕ^3 theory, i.e. Eq. (9.3.20), can be easily generalized to the case of QCD. Similarly, the generating functional of QCD can be expanded in the following form:

$$\begin{aligned} & Z\left[J, \xi, \bar{\xi}, \eta, \bar{\eta}\right] \\ & \quad = \sum_n \frac{1}{n!}\left\{\mathrm{i}\int \mathrm{d}^4 x \mathcal{L}_{\mathrm{I}}\left(\frac{\delta}{\mathrm{i}\delta J^{a\mu}}, \frac{\delta}{\mathrm{i}\delta\bar{\xi}^a}, \frac{\delta}{\mathrm{i}\delta(-\xi^a)}, \frac{\delta}{\mathrm{i}\delta\bar{\eta}}, \frac{\delta}{\mathrm{i}\delta(-\eta)}\right)\right\}^n \\ & \qquad \cdot Z_0\left[J, \xi, \bar{\xi}, \eta, \bar{\eta}\right] \end{aligned} \tag{9.3.44}$$

where Z_0 is the generating functional for free field

$$Z_0\left[J, \xi, \bar{\xi}, \eta, \bar{\eta}\right] = Z_0^{\mathrm{G}}[J]\, Z_0^{\mathrm{FP}}\left[\xi, \bar{\xi}\right] Z_0^{\mathrm{F}}\left[\eta, \bar{\eta}\right] \tag{9.3.45}$$

and Z_0^{G}, Z_0^{FP} and Z_0^{F} are the generating functionals for gauge field, ghost and quarks, respectively

$$Z_0^{\mathrm{G}} = \int [\mathrm{d}A] \exp\left\{\mathrm{i}\int \mathrm{d}^4 x\left(\mathcal{L}_0^{\mathrm{G}} + AJ\right)\right\} \tag{9.3.46}$$

$$Z_0^{\mathrm{FP}}\left[\xi, \bar{\xi}\right] = \int [\mathrm{d}\chi]\, [\mathrm{d}\bar{\chi}] \exp\left\{\mathrm{i}\int \mathrm{d}^4 x\left(\mathcal{L}_0^{\mathrm{FP}} + \bar{\chi}\xi + \bar{\xi}\chi\right)\right\} \tag{9.3.47}$$

$$Z_0^{\mathrm{F}}\left[\eta, \bar{\eta}\right] = \int [\mathrm{d}\psi]\left[\mathrm{d}\bar{\psi}\right] \exp\left\{\mathrm{i}\int \mathrm{d}^4 x\left(\mathcal{L}_0^{\mathrm{F}} + \bar{\psi}\eta + \bar{\eta}\psi\right)\right\} \tag{9.3.48}$$

In order to perform the functional differentiation in perturbation expansion of Eq. (9.3.44) easily, we need to first complete the functional integral with respect to the gauge field, ghost and quark fields in the free field generating functional. For this purpose, the integrals for the space–time coordinate in

Eqs. (9.3.46)–(9.3.48) are done by steps, and the total differential terms are eliminated, so that the free field Lagrangian can be expressed as

$$\mathcal{L}_0^{\mathrm{G}} = -\frac{1}{2}A_\mu^a K^{ab\mu\nu} A_\nu^b, \quad K_{\mu\nu}^{ab} = \delta^{ab}\left(-g_{\mu\nu}\Box + \left(1-\frac{1}{\alpha}\right)\partial_\mu\partial_\nu\right) \tag{9.3.49}$$

$$\mathcal{L}_0^{\mathrm{FP}} = -\bar{\chi}^a K^{ab}\chi^b, \quad K^{ab} = \delta^{ab}\Box \tag{9.3.50}$$

$$\mathcal{L}_0^{\mathrm{F}} = \bar{\psi}\Lambda\psi, \quad \Lambda = \mathrm{i}\gamma^\mu\partial_\mu - m \tag{9.3.51}$$

We define the inverse functions of $K_{\mu\nu}^{ab}$, K^{ab} and Λ by the following equations:

$$\int \mathrm{d}^4 z K_{\mu\lambda}^{ac}(x-z) g^{\lambda\rho} D_{\rho\nu}^{cb}(z-y) = \delta^{ab} g_{\mu\nu}\delta^4(x-y) \tag{9.3.52}$$

$$\int \mathrm{d}^4 z K^{ac}(x-z) D^{cb}(z-y) = \delta^{ab}\delta^4(x-y) \tag{9.3.53}$$

$$\int \mathrm{d}^4 z \Lambda(x-z) S(z-y) = \delta^4(x-y) \tag{9.3.54}$$

where

$$K_{\mu\nu}^{ab}(x-z) = \delta^4(x-z) K_{\mu\nu}^{ab}(z)$$

$$K^{ab}(x-z) = \delta^4(x-z) K^{ab}(z)$$

$$\Lambda(x-z) = \delta^4(x-z)\Lambda(z)$$

Solving the integral equations of (9.3.52)–(9.3.54), we can obtained these three inverse functions as

$$D_{\mu\nu}^{ab}(x-y) = \delta^{ab}\int \frac{\mathrm{d}^4 k}{(2\pi)^4}\frac{\mathrm{e}^{-\mathrm{i}k\cdot(x-y)}}{k^2+\mathrm{i}\varepsilon}\left(g_{\mu\nu} - (1-\alpha)\frac{k_\mu k_\nu}{k^2}\right) \tag{9.3.55}$$

$$D^{ab}(x-y) = \delta^{ab}\int \frac{\mathrm{d}^4 k}{(2\pi)^4}\frac{-1}{k^2+\mathrm{i}\varepsilon}\mathrm{e}^{-\mathrm{i}k\cdot(x-y)} \tag{9.3.56}$$

$$S(x-y) = \int \frac{\mathrm{d}^4 k}{(2\pi)^4}\frac{1}{\not{k} - m + \mathrm{i}\varepsilon}\mathrm{e}^{-\mathrm{i}k\cdot(x-y)} \tag{9.3.57}$$

Using these three inverse functions, we can complete the functional integrals with respect to A_μ^a, χ, $\bar{\chi}$, ψ and $\bar{\psi}$ in Eqs. (9.3.46)–(9.3.48). By omitting an insignificant integral constant, the generating functional for free fields can be obtained

$$Z_0^{\mathrm{G}}\left[J\right] = \exp\left\{\frac{\mathrm{i}}{2}\int \mathrm{d}^4x\mathrm{d}^4y J^{a\mu}(x)D^{ab}_{\mu\nu}(x-y)J^{b\nu}(y)\right\} \tag{9.3.58}$$

$$Z_0^{\mathrm{FP}}\left[\xi,\bar{\xi}\right] = \exp\left\{\mathrm{i}\int \mathrm{d}^4x\mathrm{d}^4y\bar{\xi}^a(x)D^{ab}(x-y)\xi^b(y)\right\} \tag{9.3.59}$$

$$Z_0^{\mathrm{F}}\left[\eta,\bar{\eta}\right] = \exp\left\{\mathrm{i}\int \mathrm{d}^4x\mathrm{d}^4y\bar{\eta}(x)S(x-y)\eta(y)\right\} \tag{9.3.60}$$

Substituting these free filed functional integrals into Eq. (9.3.44), we can obtain the perturbation expansion for the generating functional of QCD

$$Z\left[J,\xi,\bar{\xi},\eta,\bar{\eta}\right] = \left\{1 + \mathrm{i}\int \mathrm{d}^4x\mathcal{L}_{\mathrm{I}}\left(\frac{\delta}{\mathrm{i}\delta J^{a\mu}(x)},\ldots\right) + \cdots\right\} Z_0\left[J,\xi,\bar{\xi},\eta,\bar{\eta}\right] \tag{9.3.61}$$

Substituting Eq. (9.3.43), the interaction part of the Lagrangian of QCD $\mathcal{L}_{\mathrm{I}}$, into Eq. (9.3.61), we can calculate the generating functional order by order in perturbation expansion. For example, let us calculate the contribution of the second term in Eq. (9.3.43) to the first order in perturbative expansion, i.e. the first order of 4-gluon interaction

$$\begin{aligned} &\mathrm{i}\int \mathrm{d}^4x\mathcal{L}_{\mathrm{I}}^{4\mathrm{G}}\left(\frac{\delta}{\mathrm{i}\delta J^{a\mu}}\right)Z_0^{\mathrm{G}}\left[J\right] \\ &\quad = -\mathrm{i}\frac{g^2}{4}\int \mathrm{d}^4x f^{abe}f^{cde}\frac{\delta^4}{\delta J^{a\mu}\delta J^{b\nu}\delta J^c_\mu \delta J^d_\nu}Z_0^{\mathrm{G}}\left[J\right] \end{aligned} \tag{9.3.62}$$

where only the generating functional for gluons is considered, since other terms in free field generating functional do not contribute to this result. It is not difficult to complete the differentiation in Eq. (9.3.62). After a few steps of calculation, if only the fourth-order external source terms are retained, we can obtain

$$\begin{aligned} &-\mathrm{i}\frac{g^2}{4}f^{abe}f^{cde}\int \mathrm{d}x\mathrm{d}y_1\cdots\mathrm{d}y_4 D^{aa_1}_{\mu\mu_1}(x-y_1)D^{ba_2}_{\nu\mu_2}(x-y_2) \\ &\quad \times D^{ca_3\mu}_{\mu_3}(x-y_3)D^{ba_4\nu}_{\mu_4}(x-y_4) \\ &\quad \times J^{a_1\mu_1}(y_1)\cdots J^{a_4\mu_4}(y_4) \end{aligned} \tag{9.3.63}$$

Using this result, we can further calculate the four-gluon Green function up to the first order

$$
\begin{aligned}
&G^{ijkl}_{4\mu\nu\rho\sigma}(x_1,x_2,x_3,x_4)\\
&\quad=\langle 0|\,T\left[A^i_\mu(x_1)A^j_\nu(x_2)A^k_\rho(x_3)A^l_\sigma(x_4)\right]|0\rangle\\
&\quad=\frac{(-\mathrm{i})^4}{Z\,[0,\cdots]}\left.\frac{\delta^4 Z\,[J,\cdots]}{\delta J^{i\mu}(x_1)\delta J^{j\nu}(x_2)\delta J^{k\rho}(x_3)\delta J^{l\sigma}(x_4)}\right|_{J=0}
\end{aligned}
\tag{9.3.64}
$$

Substitute Eq. (9.3.63) into Eq. (9.3.64) to replace $Z\,[J,\cdots]$, and notice that $Z\,[0,\cdots]=1$ in the case of neglecting the integral constant when we do the functional integration to obtain Z^{G}_0, Z^{FP}_0, Z^{F}_0, i.e. Eqs. (9.3.58)–(9.3.60), then we can obtain

$$
\begin{aligned}
&G^{ijkl}_{4\mu\nu\rho\sigma}(x_1,x_2,x_3,x_4)\\
&\quad=-\mathrm{i}g^2\Big\{f^{i'j'm}f^{k'l'm}(g^{\mu'\rho'}g^{\nu'\sigma'}-g^{\mu'\sigma'}g^{\nu'\rho'})\\
&\qquad+f^{i'k'm}f^{j'l'm}(g^{\mu'\nu'}g^{\rho'\sigma'}-g^{\mu'\sigma'}g^{\nu'\rho'})+f^{i'l'm}f^{j'k'm}\\
&\qquad\times\,(g^{\mu'\nu'}g^{\rho'\sigma'}-g^{\mu'\rho'}g^{\nu'\sigma'})\Big\}\int \mathrm{d}^4x D^{i'i}_{\mu'\mu}(x-x_1)D^{j'j}_{\nu'\nu}(x-x_2)\\
&\qquad\times\,D^{k'k}_{\rho'\rho}(x-x_3)D^{l'l}_{\sigma'\sigma}(x-x_4)
\end{aligned}
\tag{9.3.65}
$$

By applying the same method, we can also calculate other Green functions, such as the three-point Green function for the three-gluon field

$$\langle 0|\,T\left[A^i_\mu(x_1)A^j_\nu(x_2)A^k_\rho(x_3)\right]|0\rangle$$

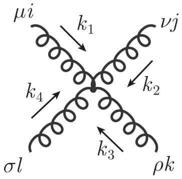
ghost–ghost–gluon fields three-point Green function

$$\langle 0|\,T\left[\bar\chi^i(x_1)\chi^j(x_2)A^k_\mu(x_3)\right]|0\rangle$$

and quark–quark–gluon fields three-point Green function

$$\langle 0|\,T\left[\bar\psi(x_1)\psi(x_2)A^i_\mu(x_3)\right]|0\rangle$$

etc. The Feynman rule can be read off from the Green functions calculated. For example, the rule of four-gluon vertex can be read off from Eq. (9.3.65)



$$
\begin{aligned}
&-\mathrm{i}g^2\left\{f^{ijm}f^{klm}(g^{\mu\rho}g^{\nu\sigma}-g^{\mu\sigma}g^{\nu\rho})\right.\\
&\quad+f^{ikm}f^{jlm}(g^{\mu\nu}g^{\rho\sigma}-g^{\mu\sigma}g^{\nu\rho})+f^{ilm}f^{jkm}\\
&\quad\left.\cdot\,(g^{\mu\nu}g^{\rho\sigma}-g^{\mu\rho}g^{\nu\sigma})\right\}\delta^4(k_1+k_2+k_3+k_4)
\end{aligned}
\tag{9.3.66}
$$

where the four-momentum conservation factor is the result of the integration of the space–time coordinate x in Eq. (9.3.65). This point can be seen directly by substituting the propagator in Eq. (9.3.55) $D^{ab}_{\mu\nu}(x-y)$ into Eq. (9.3.65). Similarly, the other Feynman rules can also be obtained after comprehensive calculations.

The Feynman rule for QCD is listed in Table 9.1, where the definitions for the functions $d_{\mu\nu}$, $V_{\mu\nu\rho}$, $W^{ijkl}_{\mu\nu\rho\sigma}$ are:

$$d_{\mu\nu}(k)=g_{\mu\nu}-(1-\alpha)\frac{k_\mu k_\nu}{k^2+\mathrm{i}\varepsilon}\tag{9.3.67}$$

$$V_{\mu\nu\rho}(k_1,k_2,k_3)=g_{\mu\nu}(k_1-k_2)_\rho+g_{\nu\rho}(k_2-k_3)_\mu+g_{\rho\mu}(k_3-k_1)_\nu\tag{9.3.68}$$

$$
\begin{aligned}
W^{ijkl}_{\mu\nu\rho\sigma}=&\left\{f^{ijm}f^{klm}(g^{\mu\rho}g^{\nu\sigma}-g^{\mu\sigma}g^{\nu\rho})\right.\\
&+f^{ikm}f^{jlm}(g^{\mu\nu}g^{\rho\sigma}-g^{\mu\sigma}g^{\nu\rho})\\
&\left.+f^{ilm}f^{jkm}\,(g^{\mu\nu}g^{\rho\sigma}-g^{\mu\rho}g^{\nu\sigma})\right\}
\end{aligned}
\tag{9.3.69}
$$

A few statements about the Feynman rules are presented below.

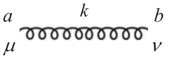
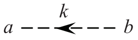
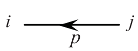
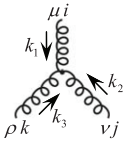
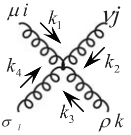
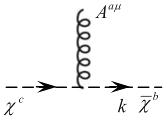
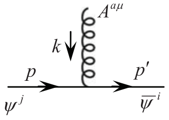
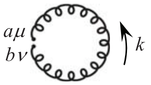
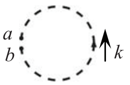
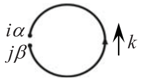
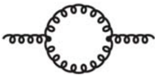
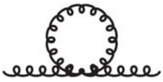
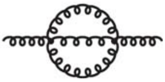
(1) For the propagator of fermion, the momentum p is aligned with the direction of the arrow of the propagator. Otherwise, the sign of the momentum in the propagator should be opposite.

(2) The free momentum for every closed loop should be integrated $\int\frac{\mathrm{d}^4k}{(2\pi)^4}$. For quark and ghost loops, an extra negative sign should be imposed. This the result of the anti-commutativity of fermions. For the loop composed of quark fields, the amplitude formed by Dirac spinor matrices in the loop should be traced.

(3) For the symmetry factor of gluon loop in diagram (a) of Table 9.1, a possibility factor 2 should be resulted by interchanging these two vertices. There are three possibilities for selecting one gluon as external line for each of the three-gluon vertex, which gives a factor 3^2. Connecting two lines of each vertex to make up inner lines gives a factor 2. Then considering the second-order factor of $1/2!$, the total symmetry factor of this loop is

$$\frac{1}{2!}\times 2\times 3^2\times 2$$

Table 9.1. The Feynman rules fro QCD.

propagator	gluon		$\delta_{ab}\frac{-id_{\mu\nu}(k)}{k^2+i\varepsilon}$
	ghost		$\delta_{ab}\frac{i}{k^2+i\varepsilon}$
	quark		$\delta_{ij}\frac{i}{\not{p}-m+i\varepsilon}$
vertex	3-gluon	 μ i k₁ k₂ k₃ ρ k ν j	$gf^{ijk}V_{\mu\nu\rho}(k_1,k_2,k_3)$
	4-gluon	 μ i k₁ νj k₂ k₄ k₃ σ l ρ k	$-ig^2W^{ijkl}_{\mu\nu\rho\sigma}$
	ghost-gluon	 $A^{a\mu}$ χ^c k $\bar{\chi}^b$	$gf^{abc}k^\mu$
	quark-gluon	 $A^{a\mu}$ k p p′ ψ^j $\bar{\psi}^i$	$igT^a_{ij}\gamma^\mu$
loop integral	gluon loop		$\int\frac{d^4k}{(2\pi)^4}\delta^{ab}g^{\mu\nu}$
	ghost loop		$-\int\frac{d^4k}{(2\pi)^4}\delta^{ab}$
	quark loop		$-\int\frac{d^4k}{(2\pi)^4}\delta^{ij}\delta^{\alpha\beta}$
	other loops		$\int\frac{d^4k}{(2\pi)^4}$
symmetry factor	 $\frac{1}{2!}$ (a)	 $\frac{1}{2!}$ (b)	 $\frac{1}{3!}$ (c)

In addition, when using the above Feynman rule to calculate this loop diagram, for each vertex there are 3! terms having already been contained, which can be seen from the number of the terms in the three-gluon vertex function V. Therefore, if calculate this loop diagram directly with the rule of the vertex, there are actually $(3!)^2$ possibilities contained, which differs from the above analysis by a factor of 1/2!. So this factor 1/2! should be compensated in the calculated result of the loop diagram. For the diagrams of (b) and (c), similar analysis is applicable. What should be noted is that there are actually 4! possibilities contained in the four-gluon vertex. In short, the symmetry factor is the difference between the calculated result of perturbation expansion and that of the direct application of the Feynman rules, so it needs to be compensated.

(4) Ghost field is the auxiliary field appearing in the quantization of the gauge field. Ghost particle is not real physical particle, so ghost particle does not appear as external line in Feynman diagrams describing physical amplitude.

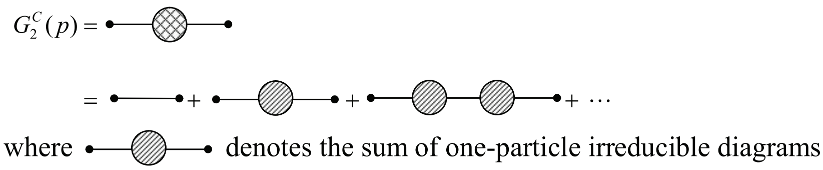
9.4. Divergence and Regularization of Loop Corrections

After establishing the Lagrangian of QCD and deriving Feynman rules for perturbation expansion, now we can calculate the amplitudes of quark and gluon scattering processes in perturbation theory. Generally, the calculation of tree level amplitude is simple, but tree level calculation cannot reflect the characteristics of QCD. Only by taking into account the corrections of loop diagrams, can we see the essentials of QCD as a non-Abelian gauge field theory. But the contribution of the loop diagram is generally divergent, and there are usually two sources for divergences. One is caused by the loop momentum going to infinity, which is called ultraviolet divergence. The other is caused as the loop momentum going to zero, which is called infrared divergence. Ultraviolet divergence can be dealt with by renormalization method, which is a method of absorbing infinity into the redefinition of the quantities such as the field, the coupling constant and particle masses, and the redefined masses and coupling constant are replaced by values of physical measurements. In this section, we introduce the regularization of ultraviolet divergence, and then to discuss the renormalization method and the treatment of ultraviolet divergence in Sec. 9.5. Meanwhile, the infrared divergence needs to be treated case by case. In some cases, the infrared divergence can be separated by factorization method, and then be absorbed into some non-perturbative physical parameters.

As an example of treating the divergence of loop diagrams in QCD, we consider the calculation of the quark self-energy diagram. Quark self-energy is part of two-point Green functions for quark fields. The connected two-point Green function (or full propagator) can be decomposed into one-particle reducible and one-particle irreducible structures. One-particle reducible diagram means that the whole Feynman diagram can be broken into two parts by cutting one particle line, while one-particle irreducible diagram means that it is impossible to split the entire Feynman diagram into two parts by disconnecting one particle line. Suppose $\Sigma_{ij}(p)$ is the connected, one-particle irreducible part of a two-point Green function, where i, j are the color indices for quarks, and p the momentum of the external quark line. Σ_{ij} includes the sum of all the one-particle irreducible Feynman diagrams. Denote the free quark propagator as $S_0 = \frac{\mathrm{i}}{\not{p}-m}$, then the connected, two-point Green function, i.e. the total propagator, can be expanded as

$$\begin{aligned} G^c_{2ij}(p) &= \int \mathrm{d}^4x \mathrm{e}^{-\mathrm{i}p\cdot x} \langle 0| T\left[\psi_i(x)\bar{\psi}_j(0)\right]|0\rangle \\ &= \frac{\mathrm{i}\delta_{ij}}{\not{p}-m} + \frac{\mathrm{i}}{\not{p}-m}\Sigma_{ij}(p)\frac{\mathrm{i}}{\not{p}-m} \\ &\quad + \frac{\mathrm{i}}{\not{p}-m}\Sigma_{il}(p)\frac{\mathrm{i}}{\not{p}-m}\Sigma_{lj}(p)\frac{\mathrm{i}}{\not{p}-m} + \cdots \\ &= S_0\delta_{ij} + S_0\Sigma_{ij}S_0 + S_0\Sigma_{il}S_0\Sigma_{lj}S_0 + \cdots \end{aligned} \tag{9.4.1}$$

Equation (9.4.1) can be expressed graphically by



Since vacuum should be color neutral, one can know $\Sigma_{ij} = \delta_{ij}\Sigma$, so Eq. (9.4.1) can be written as

$$\begin{aligned} G^c_{2ij}(p) &= \delta_{ij}\left[S_0 + S_0\Sigma S_0 + S_0\Sigma S_0\Sigma S_0 + \cdots\right] \\ &= \delta_{ij}\frac{1}{\frac{1}{S_0} - \Sigma} = \frac{\delta_{ij}}{-i(\not{p}-m) - \Sigma} \end{aligned} \tag{9.4.2}$$

The proof of Eq. (9.4.2) is in the following form. Denote

$$S = S_0 + S_0\Sigma S_0 + S_0\Sigma S_0\Sigma S_0 + \cdots \tag{9.4.3}$$

Since the right-hand side of Eq. (9.4.3) is a sum of an infinite series, extracting a factor $S_0\Sigma$ to the left from the beginning of the second term, then the sum of the residue terms is still S, i.e.

$$S = S_0 + S_0\Sigma S \tag{9.4.4}$$

From Eq. (9.4.4), we can solve

$$S = \frac{1}{\frac{1}{S_0} - \Sigma} \tag{9.4.5}$$

The amputated two-point Green function G^{tc}_{2ij} is reciprocal to the connected two-point Green function G^{c}_{2ij}. (The reason is that what being cut off is exactly the whole propagator G^{c}_{2ij} itself. Being cut once results in 1, and cut twice results in the reciprocal). So from Eq. (9.4.2), one obtain the amputated two-point Green function to be

$$G^{tc}_{2ij}(p) = \delta_{ij}\left\{-\mathrm{i}(\not{p} - m) - \Sigma\right\} \tag{9.4.6}$$

Up to the order of g^2, we can calculate Σ_{ij} using the Feynman rules obtained previously. According to the Feynman diagram in Fig. 9.3, we can write the amplitude for the quark self-energy diagram Σ_{ij} up to the order of g^2

$$\Sigma_{ij}(p) = \int \frac{\mathrm{d}^4k}{(2\pi)^4}\mathrm{i}gT^a_{il}\gamma_\mu \frac{\mathrm{i}\delta_{ln}}{\not{p} - \not{k} - m + \mathrm{i}\varepsilon}\mathrm{i}gT^b_{nj}\gamma_\nu \frac{-\mathrm{i}\delta_{ab}}{k^2 + \mathrm{i}\varepsilon}\mathrm{d}^{\mu\nu}(k) \tag{9.4.7}$$

where $d_{\mu\nu}(k) = g_{\mu\nu} - (1-\alpha)\frac{k_\mu k_\nu}{k^2+\mathrm{i}\varepsilon}$, which can be seen in Eq. (9.3.67). For SU(N) group, the generators satisfy

$$T^a_{ij}T^a_{kl} = \frac{1}{2}\left(\delta_{il}\delta_{jk} - \frac{1}{N}\delta_{ij}\delta_{kl}\right) \tag{9.4.8}$$

According to Eq. (9.4.8), we can derive

$$T^a_{ik}T^a_{kl} = (T^aT^a)_{il} = \frac{N^2-1}{2N}\delta_{il} \tag{9.4.9}$$

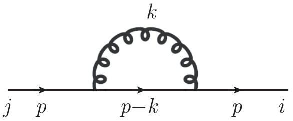


Fig. 9.3. The one-loop diagram for quark self-energy.

Equation (9.4.9) can be also written as

$$T^a_{ik}T^a_{kj} = (T^aT^a)_{ij} = C_\mathrm{F}\delta_{ij}, \quad C_\mathrm{F} = \frac{N^2-1}{2N} \tag{9.4.10}$$

Using Eq. (9.4.10), we can get

$$\Sigma_{ij}(p) = \delta_{ij}\Sigma(p) \tag{9.4.11}$$

$$\Sigma(p) = -g^2 C_\mathrm{F} \int \tfrac{\mathrm{d}^4k}{(2\pi)^4} \tfrac{\gamma_\mu(\not{p}-\not{k}+m)\gamma_\nu \mathrm{d}^{\mu\nu}(k)}{(k^2+\mathrm{i}\varepsilon)((p-k)^2-m^2+\mathrm{i}\varepsilon)} \tag{9.4.12}$$

If taking Feynman gauge, i.e. $\alpha = 1$, then $\mathrm{d}^{\mu\nu}(k) = \mathrm{g}^{\mu\nu}$, the calculation can be greatly simplified. Under the Feynman gauge

$$\Sigma(p) = -g^2 C_\mathrm{F} \int \frac{\mathrm{d}^4k}{(2\pi)^4} \frac{\gamma_\mu(\not{p}-\not{k}-m)\gamma^\mu}{(k^2+\mathrm{i}\varepsilon)\left((p-k)^2-m^2+\mathrm{i}\varepsilon\right)} \tag{9.4.13}$$

From Eq. (9.4.13), it is not difficult to find the behavior of the integration in the large momentum region

$$\Sigma(p) \sim \int \mathrm{d}^4k \frac{\not{k}}{k^2k^2} \sim \lim_{k\to\infty} k \tag{9.4.14}$$

So this integral is divergent, and the divergence comes from the contribution of the high momentum region, which is linear divergence according to this analysis. In order to make the calculation of self-energy meaningful, the first step is to make the integral convergent mathematically, whereas the actual divergence still exists as some kind of limiting behavior. Such a treatment is called regularization.

Regularization is purely a mathematical skill without a physical meaning. Therefore, in principle, there can be many ways of regularization. At present, there are usually four main kinds of ways.

(1) *Momentum cutoff method*: In this method, a maximum value Λ for momentum is taken, and the integration is done in the region lower than the maximum value Λ, so as to make the momentum loop integration convergent, and the divergence can be recovered as $\Lambda \to \infty$. Momentum truncation destroys translational and gauge invariance.

(2) *Pauli–Villars regularization*: In this method, a propagator of a particle with large mass is subtracted from the propagator in the original

theory, i.e.

$$\frac{1}{k^2 - m^2 + \mathrm{i}\varepsilon} - \frac{1}{k^2 - M^2 + \mathrm{i}\varepsilon} = \frac{m^2 - M^2}{(k^2 - m^2 + \mathrm{i}\varepsilon)(k^2 - M^2 + \mathrm{i}\varepsilon)} \xrightarrow{k \to \infty} \frac{1}{k^4} \tag{9.4.15}$$

where the behavior of the propagator changes from the original $O\left(1/k^2\right)$ to $O\left(1/k^4\right)$, so that the power of the integrand depending on the momentum greatly decreases in the high momentum region, then the integral becomes convergent. After momentum integration, taking $M^2 \to \infty$ again, the integral can recover its original value, which is divergent certainly. But the divergence has already been separated into the terms containing M^2, thus the goal of regularization is realized [15]. In Pauli–Villars regularization, the translational invariance, Lorentz invariance, and the gauge invariance for massless gauge field are all maintained. But, for massive gauge field, gauge invariance may be violated. For example, self-energy of W boson in the theory of electroweak unification cannot be properly regularized in Pauli–Villars regularization [16].

(3) *Analytic continuation method*: In this method, by changing the power of the denominator of the propagator, the power of the momentum dependence of the integrand at the high momentum region is reduced, so that the whole integral becomes convergent, i.e.

$$\frac{1}{k^2 - m^2} \to \frac{1}{(k^2 - m^2)^\alpha} \tag{9.4.16}$$

where α is a complex number with real part larger than 1, Re $\alpha > 1$. As $\alpha \to 1$, the integral recovers its original value. This method may also violate gauge invariance.

(4) *Dimensional regularization*: In this method, the integration over momentum is made convergent by varying the dimension of space–time coordinate. We can see from Eq. (9.4.14) that, if the dimension of the integral space is less than 4 instead of 4, then the integral in the ultraviolet region will tend to be convergent. Therefore, the space–time dimension is taken to be $D < 4$ in dimensional regularization, and the integration is performed in D-dimensional space. The integral is convergent in the ultraviolet limit, and it recovers to be divergent in the limit $D \to 4$ [17–20]. All symmetries, such as translation invariance, Lorentz invariance, gauge symmetry and so on, hold in dimensional

regularization. Therefore, this regularization method has been widely used in the calculation of gauge theory.

The integral in Eq. (9.4.13) will be dealt with by the dimensional normalization in the following.

At first, the 4-dimensional integral in Eq. (9.4.13) needs to be changed to D-dimensional, i.e. $\mathrm{d}^4k \to \mathrm{d}^Dk$. So, we need to make some conventions about integral in D dimension to avoid any possible ambiguity in the translation from dimension 4 to D.

The space–time index μ is changed from 0 to $D-1$ in D dimension, so the momentum should also be changed to be D-dimensional, the components of which are

$$p^\mu = (p^0, p^1, \ldots, p^{D-1}) \tag{9.4.17}$$

Thus the integral measure of the momentum space becomes D-dimensional $\mathrm{d}^4k \to \mathrm{d}^Dk$. We can either keep the factor $(2\pi)^4$ unchanged, or change it from $(2\pi)^4 \to (2\pi)^D$. Both ways are allowed as long as the limit $D \to 4$ is taken after the integration, then the result goes back to that of dimension 4. Let us do it in the following way:

$$\int \frac{\mathrm{d}^4k}{(2\pi)^4} \to \int \frac{\mathrm{d}^Dk}{(2\pi)^D} \tag{9.4.18}$$

For the space matrix element, we have

$$g^\mu_\mu = g^{\mu\nu} g_{\mu\nu} = D \tag{9.4.19}$$

The algebra relation of the Dirac γ matrices also becomes D-dimensional, i.e. γ^μ, $\mu = 0, \ldots, D-1$. In the space of dimension 4, γ matrices satisfy anti-commuting relation

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu} \tag{9.4.20}$$

To avoid the possible ambiguity, such anti-commuting relation is also kept in D-dimensional space, i.e. γ matrices are still treated as 4×4 matrices. Therefore, from Eq. (9.4.20), we can obtain

$$\mathrm{Tr}\,[\gamma_\mu \gamma_\nu] = 4g_{\mu\nu} \tag{9.4.21}$$

The matrix γ_5 do not appear in the present example, but it does appear when we apply QCD to the problem of of weak interaction processes. So it is necessary to discuss the way of treating γ_5 in D-dimensional space.

γ_5 is defined as $\gamma_5 = \mathrm{i}\gamma^0\gamma^1\gamma^2\gamma^3$ in dimension 4. A simply idea is to generalize the definition to be $\gamma_5 = \mathrm{i}\gamma^0\gamma^1\gamma^2\cdots\gamma^{D-1}$ in D-dimensional space. But rigorous analysis shows that such a simple generalization is logically problematic. Here we will not go into detail about this topic. A simple and effective way to do this is to view γ_5 as a quantity that obeys anti-commuting relation with other γ matrices, that is

$$\{\gamma_5, \gamma^\mu\} = 0, \quad \mu = 0, 1, 2, \ldots, D-1 \tag{9.4.22}$$

As long as there is no anomaly [21, 22] in problems we are dealing with, treating γ_5 in such a way will not be problematic. With these conventions, more useful equalities about γ matrices in D dimension can be listed in the following:

$$\begin{gathered}
\gamma_\mu\gamma^\mu = D, \quad \gamma_\mu\gamma^\alpha\gamma^\mu = (2-D)\gamma^\alpha, \quad \gamma_\mu\gamma_5\gamma^\mu = -D\gamma_5 \\
\gamma^\mu\gamma^\alpha\gamma^\beta\gamma_\mu = 4g^{\alpha\beta} + (D-4)\gamma^\alpha\gamma^\beta \\
\gamma^\mu\gamma^\alpha\gamma^\beta\gamma^\delta\gamma_\mu = -2\gamma^\delta\gamma^\beta\gamma^\alpha + (4-D)\gamma^\alpha\gamma^\beta\gamma^\delta \\
\mathrm{Tr}\left[\gamma^\mu\gamma^\nu\gamma^\alpha\gamma^\beta\right] = 4(g^{\mu\nu}g^{\alpha\beta} - g^{\mu\alpha}g^{\nu\beta} + g^{\mu\beta}g^{\nu\alpha})
\end{gathered} \tag{9.4.23}$$

In addition, the coupling constant of gauge field and fermions is no longer dimensionless in D-dimensional space–time. The dimension of the coupling can be derived according to the dimensionless of the action $S = \int \mathrm{d}^D x\mathcal{L}$. Let us use square bracket to denote the dimension of any quantity A, i.e. for any quantity A, its dimension is denoted as $[A]$. The Lagrangian of free fermion is $\bar{\psi}(\mathrm{i}\not{\partial} - m)\psi$. According to its dimensionless, we can obtain

$$2\,[\psi] + 1 - D = 0 \tag{9.4.24}$$

Then the dimension of fermion field in D-dimensional space–time can be derived

$$[\psi] = \frac{D-1}{2} \tag{9.4.25}$$

Similarly, according to the dimensionless of the action for gauge field, the dimension of the gauge filed A^a_μ in D-dimensional space–time is

$$[A] = \frac{D-2}{2} \tag{9.4.26}$$

The action for the interaction of gauge field and fermions is

$$\int \mathrm{d}^D x g\bar{\psi}T^a\gamma_\mu\psi A^{a\mu} \tag{9.4.27}$$

According to its dimensionless, one can obtain

$$[g] = \frac{4-D}{2} \tag{9.4.28}$$

Introducing a constant ε to denote the deviation of space–time dimension D from 4

$$D = 4 - 2\varepsilon \tag{9.4.29}$$

then the dimension of the coupling constant is

$$[g] = \varepsilon \tag{9.4.30}$$

Again introducing a new dimensionless coupling constant g_0 and a scale parameter μ with dimension of mass, then the coupling g can be expressed as

$$g = g_0 \mu^{\varepsilon} \tag{9.4.31}$$

After finishing the above preparations, we can begin to compute the quark self-energy diagram in D-dimensional space–time. In D-dimensional space–time, Eq. (9.4.13) becomes

$$\Sigma(p) = -g^2 C_{\mathrm{F}} \int \frac{\mathrm{d}^D k}{(2\pi)^D} \frac{(2-D)(\not{p} - \not{k}) + Dm}{(k^2 + \mathrm{i}\varepsilon)\left((p-k)^2 - m^2 + \mathrm{i}\varepsilon\right)} \tag{9.4.32}$$

where the first two equations in (9.4.23) are used. And then using the equation

$$\frac{1}{AB} = \int_0^1 \mathrm{d}x \frac{1}{(xA + (1-x)B)^2} \tag{9.4.33}$$

we can write Eq. (9.4.32) as

$$\begin{aligned} \Sigma(p) = &-g^2 C_{\mathrm{F}} \int \frac{\mathrm{d}^D k}{(2\pi)^D} \int_0^1 \mathrm{d}x \\ &\times \frac{(2-D)(\not{p} - \not{k}) + Dm}{\left\{x\left((p-k)^2 - m^2 + \mathrm{i}\varepsilon\right) + (1-x)(k^2 + \mathrm{i}\varepsilon)\right\}^2} \end{aligned} \tag{9.4.34}$$

Equation (9.4.33) is referred to as Feynman parameterization, and the integral over x is called Feynman parameter integral. Equation (9.4.33) can be used to calculate Feynman loop diagrams with only two inner propagators.

Meanwhile, for loop diagrams with more inner propagators, a more complicated formula for Feynman parameterization is necessary, and the more general formula for Feynman parameterization is

$$\prod_{i=1}^{n}\frac{1}{A_i^{\alpha_i}}=\frac{\Gamma(\alpha)}{\prod\limits_{i=1}^{n}\Gamma(\alpha_i)}\int_0^1\left(\prod_{i=1}^{n}\mathrm{d}x_i x_i^{\alpha_i-1}\right)\frac{\delta(1-x)}{\left(\sum_{i=1}^{n}x_iA_i\right)^{\alpha}} \tag{9.4.35}$$

where $\alpha=\sum_{i=1}^{n}\alpha_i$, $x=\sum_{i=1}^{n}x_i$. Equation (9.4.35) can be proved by recursion method, the detail for the proof is omitted here.

The denominator of Eq. (9.4.34) is rearranged, then we can obtain

$$\Sigma(p)=g^2C_{\mathrm{F}}\int_0^1\mathrm{d}x\int\frac{\mathrm{d}^Dk}{(2\pi)^D}\frac{(2-D)(\not{k}-\not{p})-Dm}{\left\{(k-xp)^2+x(1-x)p^2-xm^2+\mathrm{i}\varepsilon\right\}^2} \tag{9.4.36}$$

We have changed the order of the integrations over k and x in the above equation, since the integral is convergent as long as $D<3$, and changing the order of integration will not change the result of the integral in the case that the integral is convergent. Let us define a new integration variable

$$k'=k-xp \tag{9.4.37}$$

then

$$\begin{aligned}\Sigma(p)&=g^2C_{\mathrm{F}}\int_0^1\mathrm{d}x\int\frac{\mathrm{d}^Dk'}{(2\pi)^D}\frac{(2-D)\left(\not{k}'-(1-x)\not{p}\right)-Dm}{\left\{k'^2+x(1-x)p^2-xm^2+\mathrm{i}\varepsilon\right\}^2}\\&=g^2C_{\mathrm{F}}\int_0^1\mathrm{d}x\int\frac{\mathrm{d}^Dk'}{(2\pi)^D}\frac{-(2-D)(1-x)\not{p}-Dm}{\left\{k'^2+x(1-x)p^2-xm^2+\mathrm{i}\varepsilon\right\}^2}\end{aligned} \tag{9.4.38}$$

where the integral in the whole space is zero for an odd function, i.e. the integral for the term $\not{k}'$ in the numerator is 0. Define

$$L=-x(1-x)p^2+xm^2-\mathrm{i}\varepsilon \tag{9.4.39}$$

and moving all terms independent of momentum out of the momentum integration, we get

$$\Sigma(p)=g^2C_{\mathrm{F}}\int_0^1\mathrm{d}x\left\{-(2-D)(1-x)\not{p}-Dm\right\}\int\frac{\mathrm{d}^Dk'}{(2\pi)^D}\frac{1}{\left(k'^2-L\right)^2} \tag{9.4.40}$$

The integration for k' is what in Minkowski space. Make continuation for k'^0 to complex number. Since there are no singular points outside of the real axis for k'^0, the integration for k'^0 in the real axis can be rotated 90° to change the integration to be performed in the imaginary axis, and the result of the integral will not change. This rotation is referred to as Wick rotation, which is equivalent to defining a real momentum of K^0, and keep

$$k'^0 = \mathrm{i}K^0 \tag{9.4.41}$$

Let $\vec{k'} = \vec{K}$, then the integral measure becomes

$$\mathrm{d}^D k' = \mathrm{i}\mathrm{d}^D K \tag{9.4.42}$$

in addition

$$k'^2 = -K^2, \quad K^2 = K^{0^2} + \vec{K}^2 \tag{9.4.43}$$

After Wick rotation, the integral becomes what in Euclidean space

$$\Sigma(p) = \int \frac{\mathrm{d}^D k'}{(2\pi)^D} \frac{1}{(k'^2 - L)^2} = \int \frac{\mathrm{i}\mathrm{d}^D K}{(2\pi)^D} \frac{1}{(K^2 + L)^2} \tag{9.4.44}$$

Now let us express the D-dimensional momentum K in polar coordinates

$$\begin{aligned} K^0 &= K\cos\theta_1 \\ K^1 &= K\cos\theta_2\sin\theta_1 \\ K^2 &= K\cos\theta_3\sin\theta_2\sin\theta_1 \\ &\vdots \\ K^{D-1} &= K\sin\theta_{D-1}\sin\theta_{D-2}\cdots\sin\theta_1 \end{aligned} \tag{9.4.45}$$

where $K = (K^{0^2} + \vec{K}^2)^{1/2}$ is the magnitude of the D-dimensional momentum K. In the polar coordinate

$$\mathrm{d}^D K = K^{D-1}\mathrm{d}K\mathrm{d}\Omega_D \tag{9.4.46}$$

where

$$\mathrm{d}\Omega_D = \prod_{i=1}^{D-1} \sin^{D-1-i}\theta_i \mathrm{d}\theta_i \tag{9.4.47}$$

So the integral becomes

$$I = \frac{\mathrm{i}}{(2\pi)^D} \int \mathrm{d}\Omega_D \int_0^\infty \mathrm{d}K \frac{K^{D-1}}{(K^2 + L)^2} \tag{9.4.48}$$

Using the integration formulas

$$\int_0^\infty \mathrm{d}t \frac{t^{p-1}}{(1+t)^{p+q}} = B(p,q), \quad B(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)} \tag{9.4.49}$$

and

$$\begin{aligned}\int \mathrm{d}\Omega_D &= \int_0^\pi \mathrm{d}\theta_1 \sin^{D-2}\theta_1 \cdots \int_0^\pi \mathrm{d}\theta_{D-2} \sin\theta_{D-2} \int_0^\pi \mathrm{d}\theta_{D-1} \\ &= \frac{2\pi^{D/2}}{\Gamma\left(D/2\right)}\end{aligned} \tag{9.4.50}$$

We can obtain

$$I = \int \frac{\mathrm{d}^D k'}{(2\pi)^D} \frac{1}{\left(k'^2 - L\right)^2} = \frac{i}{(4\pi)^{D/2}} \Gamma\left(2 - D/2\right) L^{D/2-2} \tag{9.4.51}$$

For more complicated loop integrations, we need a more general formula about momentum integration in D-dimensional Minkowski space

$$\begin{aligned}\int \frac{\mathrm{d}^D k}{(2\pi)^D} \frac{(k^2)^R}{(k^2-a)^m} &= \frac{\mathrm{i}(-1)^{R-m}}{(4\pi)^{D/2}} a^{R-m+D/2} \\ &\quad \times \frac{\Gamma\left(R+D/2\right)\Gamma\left(m-R-D/2\right)}{\Gamma\left(D/2\right)\Gamma(m)}\end{aligned} \tag{9.4.52}$$

The proof of this integration formula can be done in the way given above. Equation (9.4.51) is only a special case of Eq. (9.4.52).

Substituting Eq. (9.4.51) into (9.4.40), we get

$$\begin{aligned}\Sigma(p) = g^2 C_{\mathrm{F}} \int_0^1 \mathrm{d}x \left\{-(2-D)(1-x)\, \not{p} - Dm\right\} \\ \times \frac{\mathrm{i}}{(4\pi)^{D/2}} L^{D/2-2} \Gamma\left(2 - D/2\right)\end{aligned} \tag{9.4.53}$$

Substituting Eqs. (9.4.29) and (9.4.31) into (9.4.53) and simplifying it can lead to

$$\Sigma(p) = \mathrm{i}\frac{g_0^2}{(4\pi)^2} C_{\mathrm{F}} \int_0^1 \mathrm{d}x \left\{2(1-\varepsilon)(1-x)\, \not{p} - 2(2-\varepsilon)m\right\} \left(\frac{L}{4\pi\mu^2}\right)^{-\varepsilon} \Gamma(\varepsilon) \tag{9.4.54}$$

Since the final integral result should go back to 4-dimensional space–time, i.e. $\varepsilon \to 0$, so we can expand Eq. (9.4.54) with respect to ε and only zero power term of ε needs to be kept. Using

$$\Gamma(\varepsilon)=\frac{1}{\varepsilon}-\gamma+O(\varepsilon) \tag{9.4.55}$$

$$\left(\frac{L}{4\pi\mu^2}\right)^{-\varepsilon}=1-\varepsilon\ln\frac{L}{4\pi\mu^2}+O(\varepsilon^2) \tag{9.4.56}$$

and making simplification, we can get

$$\begin{aligned}\Sigma(p)=\mathrm{i}\frac{g_0^2}{(4\pi)^2}C_{\mathrm{F}}\bigg\{&\not{p}\left(\frac{1}{\varepsilon}-\gamma+\ln4\pi-1-2\int_0^1\mathrm{d}x(1-x)\ln\frac{L}{\mu^2}\right)\\&-4m\left(\frac{1}{\varepsilon}-\gamma+\ln4\pi-\frac{1}{2}-\int_0^1\mathrm{d}x\ln\frac{L}{\mu^2}\right)\bigg\}+O(\varepsilon)\end{aligned} \tag{9.4.57}$$

Divergence in Eq. (9.4.57) can emerge again as $\varepsilon\to0$ being taken. We see that the divergence is regularized as a singular of ε, and all the other terms in this equation are finite. γ in Eq. (9.4.55) is the Euler constant, $\gamma=0.57721\ldots$.

Equation (9.4.57) is the one-loop result for the quark self-energy correction in Feynman gauge, i.e. the gauge fixing parameter is taken to be $\alpha=1$. One can also calculate the one-loop quark self-energy correction in arbitrary gauges, i.e. without taking any specific values for the gauge fixing parameter α. The only cost is that the calculation gets a little more complicated. Let us give the result for the arbitrary gauge directly in the following:

$$\begin{aligned}\Sigma(p)=\mathrm{i}\frac{g_0^2}{(4\pi)^2}C_F\bigg\{&\alpha\not{p}\left(\frac{1}{\varepsilon}-\gamma+\ln4\pi-1-2\int_0^1\mathrm{d}x(1-x)\ln\frac{L}{\mu^2}\right)\\&-m(3+\alpha)\left(\frac{1}{\varepsilon}-\gamma+\ln4\pi-\frac{1}{2}-\int_0^1\mathrm{d}x\ln\frac{L}{\mu^2}\right)\bigg\}+O(\varepsilon)\end{aligned} \tag{9.4.58}$$

9.5. The Renormalization of Divergence

The dimensional regularization is discussed in Sec. 9.4, where the divergence in loop integral is cured to be finite, and the divergence is separated into the term of $1/\varepsilon$. Here we shall discuss the renormalization method, which can eliminate the divergence emerging in higher-order corrections. The idea of renormalization is to eliminate divergence by redefining the field, the particle mass and coupling constant. The redefined physical quantity is different from the original one by a constant, which is called renormalization constant. The renormalization constant contains an infinity that can eliminates the divergent term in the loop corrections, leaving a finite value as the final result of the loop corrections. This is equivalent to absorbing infinity to the redefined field, mass and coupling constant, and then replace the redefined particle mass and coupling constant with physical measurements. With such a way to eliminate infinity, the most significant question needs to be answered is that, for a certain theory, whether the number of the infinite constants needed to eliminate infinity is finite or not. If the number of the constants is finite, this theory is a renormalizable theory, while, if it is infinite, the theory is unrenormalizable. Therefore, we need to find a criterion to judge the renormalizability of the theory. Firstly, we need a quantity to mark the degree of divergence of the loop momentum integral for a certain Feynman diagram, which is called the superficial degree of divergence of the Feynman diagram.

9.5.1. *The superficial degree of divergence*

Let us take the one-loop quark self-energy shown in Fig. 9.3 as an example. The loop-momentum integration contributed by large momentum region is

$$\Sigma(p) \sim \int \mathrm{d}^D k \frac{1}{\not{k}} \frac{1}{k^2} \tag{9.5.1}$$

where $\mathrm{d}^D k$ is the loop-momentum integral measure, $\frac{1}{\not{k}}$ is contributed by the internal fermion line, and $\frac{1}{k^2}$ the contribution of internal boson line. For the quark self-energy diagram, its divergent behavior in the large momentum limit is

$$\Sigma(p) \sim \lim_{k\to\infty} k^{D-1-2} \tag{9.5.2}$$

Therefore, the superficial degree of divergence of this Feynman diagram can be defined to be

$$d = D - 1 - 2 \tag{9.5.3}$$

where D is the contribution of one-loop integral measure, -1 and-2 the contributions of the internal lines of fermion and boson, respectively. In general, for a l-loop Feynman diagram, there are l D-dimensional momentum integral measure $\int \mathrm{d}^D k_1 \int \mathrm{d}^D k_2 \cdots \int \mathrm{d}^D k_l$. If it is restricted to consider the behavior of integral in the large momentum limit, the contribution of l D-dimensional momentum integral measures to the superficial degree of divergence is Dl. If assume that the internal line numbers of boson and fermion lines in the Ferman diagram are n_{B} and n_{F}, respectively, then their contributions to the superficial degree of divergence are $-2n_{\mathrm{B}}$ and $-n_{\mathrm{F}}$, respectively. For the quark self-energy diagram in this example, there is no momentum involved in these two vertices, so such vertices do not contribute to the superficial degree of divergence. But in general, there may be momentum in the Feynman rule for the vertex, for example, the Feynman rule for a three-gluon vertex contains momentum to the first power. If the number of momenta contained in the vertex in the Feynman diagram we consider is δ_{υ}, then the contribution of such vertex to the superficial degree of divergence is δ_{υ}. For a Feynman diagram with more such kind of vertices, the superficial degree of divergence should be the sum $\sum_{\upsilon} \delta_{\upsilon}$. Therefore, the superficial degree of divergence of a Feynman diagram is

$$d = Dl + \sum_{\upsilon} \delta_{\upsilon} - 2n_{\mathrm{B}} - n_{\mathrm{F}} \tag{9.5.4}$$

where l is the number of loops contained in the Feynman diagram, δ_{υ} the momentum factor relevant to the vertices, n_{B} the number of the internal boson lines, and n_{F} the number of internal fermion lines.

For a one-loop diagram, $d < 0$ means that the amplitude relevant to this Feynman diagram is finite, $d = 0$ means the amplitude diverges logarithmically, and $d > 0$ the amplitude diverges with higher power. However, this conclusion cannot be simply extended to multi-loop diagrams, because the conclusion of Eq. (9.5.4) is drawn by assuming that, when the loop momentum approaches infinity, all the internal line momenta approach infinity, but for some multi-loop diagrams, not all the internal line momenta approach infinity simultaneously. For example, for the two-loop diagram shown in Fig. 9.4, the momentum of the fermion line between the two one-loop diagrams is finite, which cannot be infinite, so the estimated divergence degree according to Eq. (9.5.4) does not coincide with the actual behavior of this two-loop diagram. Actually, the superficial divergence degree for the one-loop quark self-energy diagram does not coincide with its actual divergence behavior either. The superficial divergence degree estimated from Eq. (9.5.4) is $d = D - 1 - 2$. For the case of $D = 4$, the divergence degree

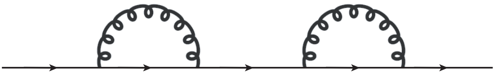


Fig. 9.4. Two-loop diagram for quark self-energy.

is $d = 1$, which is linearly divergent. However Eq. (9.4.40) resulted from the actual calculation shows that the one-loop quark self-energy diagram is logarithmically divergent. This is because Lorentz invariance causes the integral of the terms proportional to the first power of momentum in the numerator to be zero, thus reducing the actual divergence degree. Therefore, the symmetry in the scattering amplitude, such as Lorentz invariance, gauge invariance, etc. may reduce the actual degree of divergence. It is for this reason that the divergence degree given in Eq. (9.5.4) is called superficial degree of divergence.

It is also possible for higher-order diagrams that the estimated superficial degree of divergence is $d < 0$, but in fact the corresponding amplitude of this Feynman diagram diverges. For example, let us see the two-loop diagram for quark–quark scattering shown in Fig. 9.5, the superficial degree of divergence of which is $d = 2D - 4 \times 1 - 3 \times 2 = 2D - 10$, and $d = -2$ for $D = 4$. So from this result, the corresponding Feynman amplitude of this diagram should be convergent, but clearly it diverges, and the source of its divergence is from the one-loop quark self-energy diagram in the middle. Therefore, the superficial degree of divergence for higher-order diagram cannot accurately reflect the real divergence of the corresponding amplitude. For some Feynman diagrams, although their superficial degrees of divergence $d < 0$, their amplitudes diverge. The source of the divergence comes from the subdiagram, which refers to a diagram constructed by part of a Feyman diagram. The subdiagram is an untruncated Feynman diagram. For higher-order Feynman diagrams, the divergence should be decomposed in a bit more detail. The divergence is called the overall divergence, if an arbitrary internal line of the Feynman diagram is cut to form two external lines, the newly formed Feynman diagram no longer diverges. Otherwise, the divergence is called non-overall divergence.

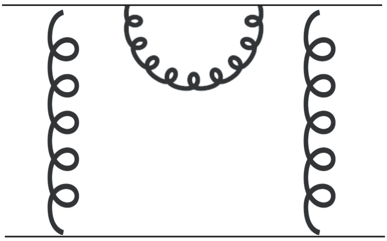


Fig. 9.5. Two-loop quark–quark scattering diagram.

Whether a Feynman diagram diverges is determined not by its superficial degree of divergence alone, but also by the superficial degrees of divergence of all its possible subdiagrams. For a Feynman diagram, in general, if the superficial degree of divergence of at least one of its subdiagram is not less than 0, the Feynman integral of the diagram diverges, unless there is some symmetry that reduces its actual divergence. Furthermore, we have the following theorem:

Convergence theorem: *For a Feynman diagram* G, *find all the possible subdiagrams* H (*including* $H = G$). *The Feynman integral corresponding to this diagram* G *is absolutely convergent if the superficial degrees of divergence for all the possible subdiagrams* H *are negative*, $d_H < 0$.

The proof of this theorem is beyond the scope of this book and will not be given here. The interested readers may refer to the original literatures [23–26].

9.5.2. *The renormalizability of a theory*

We can determine whether a theory of interaction is renormalizable with the superficial degree of divergence. Let us consider a simple case at first, i.e. the case that there is only one interaction term $\mathcal{L}_{\mathrm{I}}$ in the Lagrangian $\mathcal{L}$ of a theory of interaction.

The superficial degree of divergence given in Eq. (9.5.4) depends on the number of internal fermion and boson lines of a specific Feynman diagram. This is beneficial to judging the divergence of a specific Feynman diagram. But it is not convenient to use it to judge whether a theory of interaction is renormalizable. Therefore, we wish to express the superficial degree of divergence directly in terms of the characteristic quantities in the theory of interaction, that is, how many boson and fermion lines there are in the interaction terms, and what the momentum factor (the number of space–time derivatives) is.

Suppose b is the number of boson fields in the interaction term $\mathcal{L}_{\mathrm{I}}$, f the number of fermion fields, and δ the number of space–time derivatives in $\mathcal{L}_{\mathrm{I}}$. For a specific Feynman diagram G, suppose n is the number of vertices contained in G, N_{B} the number of external boson lines in the Feynman diagram G, and N_{F} the number of external fermion lines.

For any Feynman diagrams, an internal line is connected to the vertices twice, while an external line is connected to the vertices only once. Therefore, for a Feynman diagram G with n vertices, the numbers of boson and fermion lines should satisfy the following relation, respectively

$$2n_{\mathrm{B}} + N_{\mathrm{B}} = nb \tag{9.5.5}$$

$$2n_{\mathrm{F}} + N_{\mathrm{F}} = nf \tag{9.5.6}$$

where n_{B} and n_{F} are the numbers of internal boson and fermion lines, respectively. There are l free momentum integrals for an l-loop Feynman diagrams, but actually each internal line corresponds to one 4-momentum integral, so there should be $n_{\mathrm{B}}+n_{\mathrm{F}}$ integrals originally. There is one δ function for each vertex, which comes from momentum conservation. There are n δ functions for n vertices, which amounts to n restriction conditions. However, there should be one overall momentum conservation among these n conditions. The overall momentum conservation have nothing to do with the degree of freedom of internal momenta. Therefore, there are totally $n-1$ restrictions on the number of the free momenta. Finally, there are $n_{\mathrm{B}} + n_{\mathrm{F}} - (n-1)$ free momentum integrals, so we have the following relation:

$$l = n_{\mathrm{B}} + n_{\mathrm{F}} - n + 1 \tag{9.5.7}$$

In addition, the sum of momentum factors $\sum_{v} \delta_{v}$ in Eq. (9.5.4) is equal to the total number of momenta among n vertices

$$\sum_{v} \delta_{v} = n\delta \tag{9.5.8}$$

Using Eqs. (9.5.5) and (9.5.6) to eliminate the numbers of internal lines in Eq. (9.5.7) results in

$$l = \left(\frac{b+f}{2} - 1\right) n - \frac{N_{\mathrm{B}} + N_{\mathrm{F}}}{2} + 1 \tag{9.5.9}$$

Substituting Eq. (9.5.9) into Eq. (9.5.4), using Eqs. (9.5.5) and (9.5.6) again to eliminate the internal line numbers n_{B} and n_{F}, we obtain the superficial degree of divergence

$$d = rn - \frac{D-2}{2} N_{\mathrm{B}} - \frac{D-1}{2} N_{\mathrm{F}} + D \tag{9.5.10}$$

where r is

$$r = \frac{D-2}{2} b + \frac{D-1}{2} f + \delta - D \tag{9.5.11}$$

r only depends on the interaction part of the Lagrangian $\mathcal{L}_{\mathrm{I}}$, which is referred to as divergence index of the interaction term $\mathcal{L}_{\mathrm{I}}$. The superficial degree of divergence given in Eq. (9.5.10) is only related to r and the external line number of Feynman diagram. For a specific scattering process, the external line number of the Feynman diagram is always fixed, which has nothing to do with the order of Feynman diagram, so the expression of superficial divergence degree given in Eq. (9.5.10) is more beneficial to the analysis of the renormalizability of a theory.

There is a more explicit meaning for r in Eq. (9.5.11), which is the negative value of the dimension of the coupling constant. The interaction part of the interaction Lagrangian can be formally written as

$$\mathcal{L}_{\mathrm{I}} \sim g(\partial)^{\delta}(\phi)^{b}(\psi)^{f} \tag{9.5.12}$$

Since the action $S_{\mathrm{I}} = \int \mathrm{d}^D x \mathcal{L}_{\mathrm{I}}$ is dimensionless, we have

$$0 = -D + [g] + \delta + \frac{D-2}{2}b + \frac{D-1}{2}f \tag{9.5.13}$$

From Eq. (9.5.13), we get

$$[g] = -\left(\frac{D-2}{2}b + \frac{D-1}{2}f + \delta - D\right) \tag{9.5.14}$$

That is $[g] = -r$.

Equation (9.5.10) can be easily extended to the case with more interaction terms. Suppose that the interaction part of a theory consists of several terms, $\mathcal{L}_{\mathrm{I}} = \sum_i \mathcal{L}_i$, where the coupling constant associated with $\mathcal{L}_i$ is g_i, the numbers of boson and fermion fields are b_i and f_i, respectively, and the number of derivatives is δ_i. If the number of vertex of $\mathcal{L}_i$ is n_i, then the superficial degree of divergence of Eq. (9.5.10) should be extended to be

$$d = \sum_i r_i n_i - \frac{D-2}{2}N_{\mathrm{B}} - \frac{D-1}{2}N_{\mathrm{F}} + D \tag{9.5.15}$$

where

$$r_i = \frac{D-2}{2}b_i + \frac{D-1}{2}f_i + \delta_i - D \tag{9.5.16}$$

We can also replace r_i in Eq. (9.5.15) to $-[g_i]$, i.e.

$$d = -\sum n_i [g_i] - \frac{D-2}{2}N_{\mathrm{B}} - \frac{D-1}{2}N_{\mathrm{F}} + D \tag{9.5.17}$$

The second and third terms in Eq. (9.5.17) are also explicitly meaningful. They are the negative values of the total dimension of all the external boson lines and fermion lines of the Feynman diagram G, respectively. Therefore, Eq. (9.5.17) shows that the superficial degree of divergence of Feynman diagram G is equal to the dimension D minus all dimensions of the coupling constants in the vertices, and minus the total dimensions of all the external fields.

If at least one divergence index is larger than 0, $r_i > 0$, in all interaction terms $\mathcal{L}_i$ in a theory, then the number of divergence types in the Feynman diagram increases infinitely as the perturbation order increases. It is impossible to eliminate all the divergences with finite number of constants in such kind of a theory. Such theory is not renormalizable. On the other hand, if all $r_i \leq 0$ in a theory, then the number of divergence types are finite for any order of Feynman diagrams. Such kind of theory is renormalizable. The theory with all $r_i = 0$ is referred to as renormalizable theory, and with all $r_i < 0$ the super renormalizable theory.

In QCD, for all the four interaction terms $r_i \propto D - 4$, so QCD is renormalizable in dimension 4. The external lines of ghost are treated as boson external lines when calculating the Feynman diagrams in QCD,[a] i.e.

$$N_{\mathrm{B}} = N_{\mathrm{G}} + N_{\mathrm{FP}} \tag{9.5.18}$$

where N_{G} is the number of external lines of gluon, N_{FP} the number of Faddeev–Popov ghost external lines. In addition, the momentum factor k_μ only appears on the side of ghost line flowing outward from the vertex in the Feynman rule of gluon–ghost interaction vertex. There are always momentum k_μ corresponding to half number of external ghost lines which do not contribute to the divergence index, i.e. the factor $\sum_v \delta_v$ in Eq. (9.5.4). Therefore, it is necessary to subtract $\frac{1}{2}N_{\mathrm{FP}}$ from Eq. (9.5.4) or (9.5.15) and Eq. (9.5.17). Including ghost fields, the superficial degree of divergence of Feynman diagram in QCD becomes

$$d = \sum_i r_i n_i - \frac{D-2}{2} N_{\mathrm{B}} - \frac{D-1}{2} N_{\mathrm{F}} + D - \frac{1}{2} N_{\mathrm{FP}} \tag{9.5.19}$$

Substituting Eq. (9.5.18) into Eq. (9.5.19), one can obtain

$$d = \sum_i r_i n_i - \frac{D-2}{2} N_{\mathrm{G}} - \frac{D-1}{2} (N_{\mathrm{FP}} + N_{\mathrm{F}}) + D \tag{9.5.20}$$

---

[a] Ghost field particles are not physical particles and cannot appear as external lines in a Feynman diagram, which refers to the total Feynman diagram of physical scattering, but the Feynman diagram discussed here includes subdiagrams of a Feynman diagram of physical process.

For the case $D = 4$, $r_i = 0$, so

$$d = 4 - N_{\mathrm{G}} - \frac{3}{2}(N_{\mathrm{FP}} + N_{\mathrm{F}}) \tag{9.5.21}$$

According to Eq. (9.5.21), one can figure out that the number of superficially divergent Feynman diagrams in QCD is finite. Considering $d \geq 0$, one can find only several cases match the condition. Dropping the cases of odd number of fermion lines and ghost lines, we have $(N_{\mathrm{G}}, N_{\mathrm{FP}}, N_{\mathrm{F}}) = (0,0,0)$, $(1,0,0)$, $(2,0,0)$, $(3,0,0)$, $(4,0,0)$, $(0,2,0)$, $(1,2,0)$, $(0,0,2)$, $(1,0,2)$. Among these nine cases, $(0,0,0)$ belongs to vacuum diagram, which has been absorbed into the normalization of the generating functional, and the case of $(1,0,0)$ is removed because of Lorentz invariance. Therefore, only seven types of divergent diagrams left, which are shown in Fig. 9.6.

9.5.3. *The renormalization of QCD*

Now we begin to discuss the problem of eliminating the divergence that occurs when computing higher-order corrections in QCD, i.e. the renormalization of QCD. The Lagrangian of QCD has been given in

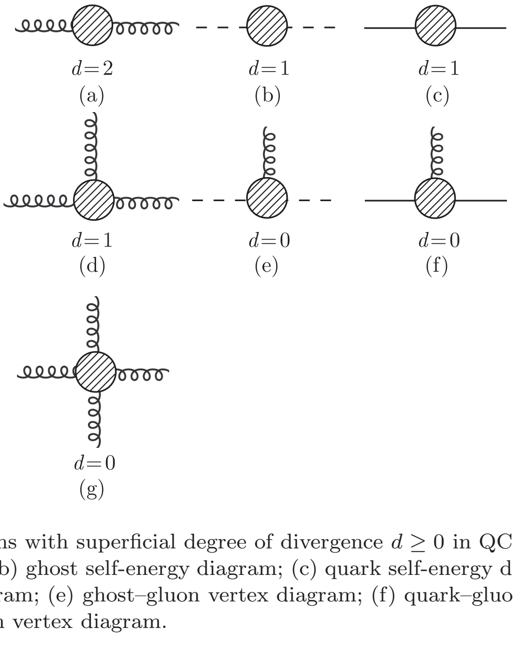


Fig. 9.6. Diagrams with superficial degree of divergence $d \geq 0$ in QCD. (a) Gluon self-energy diagram; (b) ghost self-energy diagram; (c) quark self-energy diagram; (d) three-gluon vertex diagram; (e) ghost–gluon vertex diagram; (f) quark–gluon vertex diagram; and (g) four-gluon vertex diagram.

Eq. (9.3.6), i.e.

$$\begin{aligned}\mathcal{L} = &-\frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} - \frac{1}{2\alpha}(\partial^\mu A^a_\mu)^2 + (\partial^\mu \bar{\chi}^a)D^{ab}_\mu \chi^b \\ &+ \bar{\psi}^i(\mathrm{i}\gamma^\mu D^{ij}_\mu - m\delta^{ij})\psi^j \end{aligned} \tag{9.5.22}$$

where the covariant derivatives D^{ij}_μ and D^{ab}_μ are in the basis and adjoint representations of color SU(3) group, respectively,

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + gf^{abc}A^b_\mu A^c_\nu \tag{9.5.23}$$

$$D^{ij}_\mu = \delta^{ij}\partial_\mu - \mathrm{i}gT^c_{ij}A^c_\mu \tag{9.5.24}$$

$$D^{ab}_\mu = \delta^{ab}\partial_\mu - \mathrm{i}gT^c_{ab}A^c_\mu, \quad T^c_{ab} = -\mathrm{i}f^{cab} \tag{9.5.25}$$

The Lagrangian of QCD can be separated into free part $\mathcal{L}_0$ and interaction part $\mathcal{L}_\mathrm{I}$,

$$\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_\mathrm{I} \tag{9.5.26}$$

$$\begin{aligned}\mathcal{L}_0 = &-\frac{1}{4}(\partial_\mu A^a_\nu - \partial_\nu A^a_\mu)(\partial^\mu A^{a\nu} - \partial^\nu A^{a\mu}) - \frac{1}{2\alpha}(\partial^\mu A^a_\mu)^2 \\ &+ (\partial^\mu \bar{\chi}^a)\partial_\mu \chi^a + \bar{\psi}^i(\mathrm{i}\gamma^\mu\partial_\mu - m)\psi^i \end{aligned} \tag{9.5.27}$$

$$\begin{aligned}\mathcal{L}_\mathrm{I} = &-\frac{1}{2}gf^{abc}(\partial_\mu A^a_\nu - \partial_\nu A^a_\mu)A^{b\mu}A^{c\nu} - \frac{1}{4}g^2 f^{abe}f^{cde}A^a_\mu A^b_\nu A^{c\mu}A^{d\nu} \\ &- gf^{abc}A^c_\mu(\partial^\mu \bar{\chi}^a)\chi^b + g\bar{\psi}^i T^a_{ij}\gamma^\mu\psi^j A^a_\mu \end{aligned} \tag{9.5.28}$$

There are totally four terms in the interaction part, which are three-gluon, four-gluon interactions, ghost–gluon interaction and quark–gluon interaction. Their divergence indices are: three-gluon interaction $r_1 = \frac{D-4}{2}$, four-gluon interaction $r_2 = D - 4$, gluon–ghost interaction $r_3 = \frac{D-4}{2}$, quark–gluon interaction $r_4 = \frac{D-4}{2}$, respectively. QCD is a renormalizable theory in 4-dimensional space–time according to these divergence indices, i.e. the infinities appearing in the higher-order corrections of QCD can be absorbed by introducing finite number of infinite constants.

The constants used to cancel the divergences are introduced by redefining the fields, the particle masses and coupling constant. The quarks, gluons and ghost fields are redefined as renormalized fields, and the previously discussed unrenormalized fields are called bare fields

$$\psi = Z_2^{1/2}\psi_r, \quad A^a_\mu = Z_3^{1/2}A^a_{r\mu}, \quad \chi^a = \tilde{Z}_3^{1/2}\chi^a_r, \quad \bar{\chi}^a = \tilde{Z}_3^{1/2}\bar{\chi}^a_r \tag{9.5.29}$$

The coupling constant, quark mass and gauge fixing parameter are also redefined to renormalized coupling constant, mass and gauge fixing parameter, while the unrenormalized quantities are called bare quantities

$$g = Z_g g_r, \quad m = Z_m m_r, \quad \alpha = Z_3 \alpha_r \tag{9.5.30}$$

The above Z_2, Z_3, $\tilde{Z}_3$ are the renormalization constants of quark field, gluon field and ghost field, respectively. Meanwhile Z_g and Z_m are the renormalization constants of the coupling and quark masses, respectively. The renormalization constant of gauge fixing parameter is the same as that of gluon field, which keeps the gauge fixing term $\frac{1}{2\alpha}(\partial^\mu A^a_\mu)^2$ remaining unchanged after renormalization.

Substituting Eqs. (9.5.29) and (9.5.30) into the Lagrangian of QCD, Eq. (9.5.22), we can obtain the Lagrangian in terms of renormalized fields, masses and coupling constant

$$\mathcal{L} = \mathcal{L}_{r0} + \mathcal{L}_{r\mathrm{I}} + \mathcal{L}_c \tag{9.5.31}$$

where $\mathcal{L}_{r0}$ and $\mathcal{L}_{r\mathrm{I}}$ are the renormalized free and interaction parts of the Lagrangian, which are formally equal to $\mathcal{L}_0$ and $\mathcal{L}_\mathrm{I}$ with the field, masses, coupling constant and gauge fixing parameter replaced with their relevant renormalized quantities, while $\mathcal{L}_c$ is referred to as counter term, which equals

$$\begin{aligned}
\mathcal{L}_c = {} & -(Z_3 - 1)\frac{1}{4}(\partial_\mu A^a_{r\nu} - \partial_\nu A^a_{r\mu})(\partial^\mu A^{a\nu}_r - \partial^\nu A^{a\mu}_r) \\
& + (\tilde{Z}_3 - 1)(\partial^\mu \bar{\chi}^a_r)(\partial_\mu \chi^a_r) + (Z_2 - 1)\bar{\psi}^i_r \mathrm{i}\gamma^\mu \partial_\mu \psi_r \\
& - (Z_2 Z_m - 1) m_r \bar{\psi}^i_r \psi^i_r \\
& - (Z_g Z_3^{3/2} - 1)\frac{1}{2} g_r f^{abc}(\partial_\mu A^a_{r\nu} - \partial_\nu A^a_{r\mu}) A^{b\mu}_r A^{c\nu}_r \\
& - (Z_g^2 Z_3^2 - 1)\frac{1}{4} g_r^2 f^{abe} f^{cde} A^a_{r\mu} A^b_{r\nu} A^{c\mu}_r A^{d\nu}_r \\
& - (Z_g \tilde{Z}_3 Z_3^{1/2} - 1) g_r f^{abc} A^c_{r\mu} (\partial^\mu \bar{\chi}^a_r) \chi^b_r \\
& + (Z_g Z_2 Z_3^{1/2} - 1) g_r \bar{\psi}^i_r T^a_{ij} \gamma^\mu \psi^j_r A^a_{r\mu}
\end{aligned} \tag{9.5.32}$$

Some combinations of these constants in the above equation can be defined as new renormalization constants, which can make the expression more concise. Thus one can define

$$\begin{aligned} Z_1 &\equiv Z_g Z_3^{3/2}, \qquad & Z_4 &\equiv Z_g^2 Z_3^2 \\ \tilde{Z}_1 &\equiv Z_g \tilde{Z}_3 Z_3^{1/2}, \qquad & Z_{1\mathrm{F}} &\equiv Z_g Z_2 Z_3^{1/2} \end{aligned} \tag{9.5.33}$$

So the counter term can reexpressed as

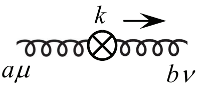
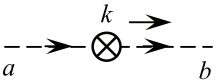
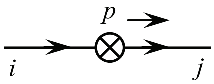
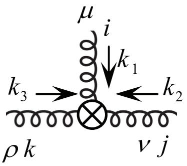
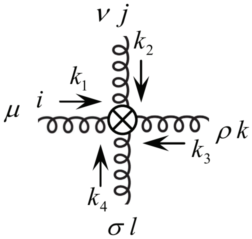
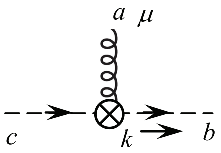
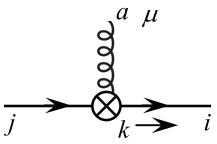
$$\begin{aligned} \mathcal{L}_c = &-(Z_3-1)\frac{1}{4}(\partial_\mu A^a_{r\nu} - \partial_\nu A^a_{r\mu})(\partial^\mu A^{a\nu}_r - \partial^\nu A^{a\mu}_r) \\ &+ (\tilde{Z}_3 - 1)(\partial^\mu \bar{\chi}^a_r)(\partial_\mu \chi^a_r) + (Z_2 - 1)\bar{\psi}^i_r \mathrm{i}\gamma^\mu \partial_\mu \psi_r \\ &- (Z_2 Z_m - 1) m_r \bar{\psi}^i_r \psi^i_r \\ &- (Z_1 - 1)\frac{1}{2} g_r f^{abc}(\partial_\mu A^a_{r\nu} - \partial_\nu A^a_{r\mu}) A^{b\mu}_r A^{c\nu}_r \\ &- (Z_4 - 1)\frac{1}{4} g_r^2 f^{abe} f^{cde} A^a_{r\mu} A^b_{r\nu} A^{c\mu}_r A^{d\nu}_r \\ &- (\tilde{Z}_1 - 1) g_r f^{abc} A^c_{r\mu}(\partial^\mu \bar{\chi}^a_r)\chi^b_r \\ &+ (Z_{1\mathrm{F}} - 1) g_r \bar{\psi}^i_r T^a_{ij} \gamma^\mu \psi^j_r A^a_{r\mu} \end{aligned} \tag{9.5.34}$$

Using the counter terms in Eq. (9.5.34), and according to the perturbative expansion procedure, we can derive the Feynman rules of the counter terms, which are given in Table 9.2

There are infinities contained in the renormalization constants. In QCD the calculation of the higher-order correction is carried on by adjusting the renormalization constants to cancel the divergence of perturbation correction order by order. This is equivalent to absorbing the divergence of the loop integral into the redefinition of the quark mass and coupling constant of the fields. In the process of renormalization, divergence in the higher-order perturbative corrections are always canceled by the counter term resulted from the lower-order expansion. The renormalization constant is also a sum of series of order term. For example, the divergence of one-loop diagram correction is canceled by the counter term of tree-level diagram, while the result of the counter term calculated up to one-loop order is used to remove the divergence of two-loop result in the perturbative expansion, and so on.

According to Eqs. (9.5.33) and (9.5.34), there are four kinds of ways to determine the renormalization constant Z_g, i.e. using the last four terms in $\mathcal{L}_c$. A significant question is whether the results of Z_g determined in these four ways are equal. If Z_g determined in these ways are all equal, according

Table 9.2. The Feynman rules for counter terms.

	$i(Z_3-1)\delta_{ab}(k_\mu k_\nu - k^2 g_{\mu\nu})$	Counter term for gluon
	$i(\widetilde{Z}_3-1)\delta_{ab}k^2$	Counter term for ghost
	$i[(Z_2-1)\not{p}-(Z_2Z_m-1)m_r]\delta_{ij}$	Counter term for quark
 μ i k_1 k_3 k_2 $\rho\, k$ $\nu\, j$	$(Z_1-1)g_r f^{ijk}V_{\mu\nu\rho}(k_1,k_2,k_3)$	Counter term for three-gluon
 $\nu\, j$ k_2 k_1 μ i $\rho\, k$ k_3 k_4 $\sigma\, l$	$(Z_4-1)(-i)g_r^2 W^{ijkl}_{\mu\nu\rho\sigma}$	Counter term for Four-gluon
	$(\widetilde{Z}_1-1)g_r f^{abc}k^\mu$	Counter term for ghost–gluon
	$i(Z_{1\mathrm{F}}-1)g_r T^a_{ij}\gamma_\mu$	Counter term for quark–gluon

to Eq. (9.5.33), Z_1, Z_4, $\tilde{Z}_1$ and $Z_{1\mathrm{F}}$ should satisfy the following equation

$$\frac{Z_1}{Z_3}=\frac{\tilde{Z}_1}{\tilde{Z}_3}=\frac{Z_{1\mathrm{F}}}{Z_2}=\frac{Z_4}{Z_1} \tag{9.5.35}$$

Equation (9.5.35) is called the Slavnov–Taylor identity in the narrow sense. We'll see in a moment, to the one-loop approximation, this equation is indeed satisfied. This is not a coincidence. It is the result of a symmetry that the Lagrangian of QCD still possesses after quantization, which is called BRST symmetry. We will discuss BRST symmetry in the next section. The generalized Ward–Takahashi identity can be derived from BRST symmetry. This identity restricts the relation between some Green functions. It can be proved that Eq. (9.5.35) can be derived by BRST symmetry.

Now, as an example, let us see how the divergence is eliminated by the selection of Z_2 and Z_m in the calculation of the quark self-energy one-loop correction. In addition to the contribution of Feynman diagram in Fig. 9.3, the contribution of quark counter term should also be considered in the calculation up to one-loop level, as shown in Fig. 9.7.

The renormalized one-loop self-energy part is denoted by $\Sigma^{\mathrm{R}}_{ij}(p)$, which is the sum of the contributions of the one-loop self-energy diagram $\Sigma_{ij}(p)$ and the counter term

$$\Sigma^{\mathrm{R}}_{ij}(p)=\Sigma_{ij}(p)+\Sigma^{\mathrm{c}}_{ij}(p) \tag{9.5.36}$$

where $\Sigma^{\mathrm{c}}_{ij}(p)$ is the contribution of the counter term. According to the Feynman rules

$$\Sigma^{\mathrm{c}}_{ij}(p)=\mathrm{i}\left[(Z_2-1)\ \not{p}-(Z_2Z_m-1)m_r\right]\delta_{ij} \tag{9.5.37}$$

With $\Sigma_{ij}(p)=\delta_{ij}\Sigma(p)$, which can be seen in Eq. (9.4.11), $Z^{\mathrm{R}}_{ij}(p)$ can be written as

$$\Sigma^{\mathrm{R}}_{ij}(p)=\delta_{ij}\Sigma^{\mathrm{R}}(p) \tag{9.5.38}$$

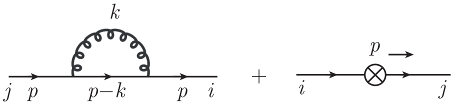


Fig. 9.7. The quark self-energy diagram and the counter term.

With the above three equations considered, the renormalized one-loop self-energy contribution is obtained

$$\Sigma^{\mathrm{R}}(p) = \Sigma(p) + \mathrm{i}\left[(Z_2 - 1)\, \not{p} - (Z_2 Z_m - 1)m_r\right] \tag{9.5.39}$$

Substituting the result of one-loop self-energy part calculated in Sec. 9.4, i.e. Eq. (9.4.57), into the above equation, we can obtain

$$\begin{aligned}\Sigma^{\mathrm{R}}(p) = \mathrm{i}\frac{g_r^2}{(4\pi)^2}C_{\mathrm{F}}\Bigg\{ &\not{p}\left(\frac{1}{\varepsilon} - \gamma + \ln 4\pi - 1 - 2\int_0^1 \mathrm{d}x(1-x)\ln\frac{L}{\mu^2}\right) \\ &- 4m_r\left(\frac{1}{\varepsilon} - \gamma + \ln 4\pi - \frac{1}{2} - \int_0^1 \mathrm{d}x \ln\frac{L}{\mu^2}\right)\Bigg\} + O(\varepsilon) \\ &+ \mathrm{i}\left[(Z_2 - 1)\, \not{p} - (Z_2 Z_m - 1)m_r\right]\end{aligned} \tag{9.5.40}$$

When substituting Eq. (9.4.57), we have made the replacement for the mass term in Eq. (9.4.57) $m \to m_r$, and dropped the non-singular terms of ε. Compare each term in Eq. (9.5.40), and select the renormalization constants Z_2 and $Z_2 Z_m$ to make them to be equal to the divergent parts of $\not{p}$ and m_r in magnitude and opposite in sign, then the divergences can be canceled and the finite one-loop correction can be obtained. However, a significant question is how to clearly define the divergent term in $\Sigma(p)$, because adding any finite term on a divergent one is also a divergent term. This will lead to the arbitrariness in defining the divergence. In order to remove the arbitrariness, we must define the divergent part explicitly, i.e. define which part in the Green function will be subtracted together with the infinity. The way to determine which terms are subtracted with infinity together is called the renormalization scheme. The selection of the renormalization scheme is not unique, which depends on the way of determining the finite terms subtracted with $1/\varepsilon$ together. Below we shall introduce several subtraction schemes which are frequently used in the literatures.

9.5.3.1. *On-shell subtraction*

Selecting Z_2 and Z_m, so that as $\not{p} \sim m_r$, the whole propagator of quark is in the form of free quark propagator, i.e.

$$G_{2ij}^{\mathrm{R}}(p) \sim \frac{\mathrm{i}\delta_{ij}}{\not{p} - m}, \quad \text{as} \quad \not{p} \sim m \tag{9.5.41}$$

After renormalization, the expression of the propagator of Eq. (9.4.2) becomes

$$G^{\mathrm{R}}_{2ij}(p) = \frac{\delta_{ij}}{-\mathrm{i}(\not{p} - m_r) - \Sigma^{\mathrm{R}}(p)} \tag{9.5.42}$$

This is equivalent to replace m and Σ in Eq. (9.4.2) as $m \to m_r$, $\Sigma \to \Sigma^{\mathrm{R}}$. To make Eq. (9.5.42) tend to Eq. (9.5.41) as $\not{p} \sim m_r$, we only need that, as $\not{p} \sim m_r$,

$$\Sigma^{\mathrm{R}}(p) \sim 0 \tag{9.5.43}$$

As $\not{p} \sim m_r$

$$\begin{aligned} L &= -x(1-x)p^2 + xm_r^2 - \mathrm{i}\varepsilon \\ &\sim m_r^2 x^2 - \mathrm{i}\varepsilon \end{aligned} \tag{9.5.44}$$

Substituting the above result of L into Eq. (9.5.40), we obtain that, as $\not{p} \sim m_r$,

$$\begin{aligned} \Sigma^{\mathrm{R}}(p) \sim \mathrm{i}\frac{g_r^2}{(4\pi)^2} C_{\mathrm{F}} \Bigg\{ \not{p} \left(\frac{1}{\varepsilon} - \gamma + \ln 4\pi + 2 - \ln \frac{m_r^2}{\mu^2} \right) \\ - 4m_r \left(\frac{1}{\varepsilon} - \gamma + \ln 4\pi + \frac{3}{2} - \ln \frac{m_r^2}{\mu^2} \right) \Bigg\} \\ + \mathrm{i}\left[(Z_2 - 1)\not{p} - (Z_2 Z_m - 1) m_r \right] \end{aligned} \tag{9.5.45}$$

Requiring Eq. (9.4.45) tend to 0, one can obtain

$$Z_2 = 1 - \frac{g_r^2}{(4\pi)^2} C_{\mathrm{F}} \left(\frac{1}{\varepsilon} - \gamma + \ln 4\pi + 2 - \ln \frac{m_r^2}{\mu^2} \right) \tag{9.5.46}$$

$$Z_2 Z_m = 1 - \frac{4g_r^2}{(4\pi)^2} C_{\mathrm{F}} \left(\frac{1}{\varepsilon} - \gamma + \ln 4\pi + \frac{3}{2} - \ln \frac{m_r^2}{\mu^2} \right) \tag{9.5.47}$$

Dividing Z_2 on the left-hand side of Eq. (9.5.47) to the right, substituting the result of Eq. (9.5.46), and expanding the result with respect to g_r, one obtains

$$\begin{aligned}Z_m &= 1 - \frac{4g_r^2}{(4\pi)^2}C_{\mathrm{F}}\left(\frac{1}{\varepsilon} - \gamma + \ln 4\pi + \frac{3}{2} - \ln\frac{m_r^2}{\mu^2}\right)\\ &\quad + \frac{g_r^2}{(4\pi)^2}C_{\mathrm{F}}\left(\frac{1}{\varepsilon} - \gamma + \ln 4\pi + 2 - \ln\frac{m_r^2}{\mu^2}\right) + O(g_r^4)\\ &= 1 - \frac{3g_r^2}{(4\pi)^2}C_{\mathrm{F}}\left(\frac{1}{\varepsilon} - \gamma + \ln 4\pi + \frac{4}{3} - \ln\frac{m_r^2}{\mu^2}\right) + O(g_r^4) \qquad (9.5.48)\end{aligned}$$

9.5.3.2. *Off-shell subtraction*

For the case of quark mass $m_r \sim 0$, require that the propagator tends to be free quark propagator as p^2 is a value less than zero, $p^2 = -\lambda^2 < 0$,

$$G_{2ij}^{\mathrm{R}}(p) \sim \frac{\mathrm{i}\delta_{ij}}{\not p}, \quad \text{as} \quad p^2 \sim -\lambda^2 \qquad (9.5.49)$$

Substituting $m_r = 0$ into Eq. (9.5.40), and noticing that $L = -x(1-x)p^2 + m_r^2 x - \mathrm{i}\varepsilon$, one can obtain for the case of zero mass

$$\Sigma^{\mathrm{R}}(p) = \mathrm{i}\frac{g_r^2}{(4\pi)^2}C_{\mathrm{F}}\ \not p\left(\frac{1}{\varepsilon} - \gamma + \ln 4\pi + 1 - \ln\frac{-p^2}{\mu^2}\right) + \mathrm{i}(Z_2 - 1)\ \not p \qquad (9.5.50)$$

With $m_r = 0$, from Eq. (9.5.42), one can get

$$G_{2ij}^{\mathrm{R}}(p) = \frac{\delta_{ij}}{-\mathrm{i}\ \not p - \Sigma^{\mathrm{R}}(p)} \qquad (9.5.51)$$

Therefore, the requirement of Eq. (9.5.49) is equivalent to $\Sigma^{\mathrm{R}}(p) \sim 0$ as $p^2 \sim -\lambda^2$. So, from Eq. (9.5.50), we get

$$Z_2 = 1 - \frac{g_r^2}{(4\pi)^2}C_{\mathrm{F}}\left(\frac{1}{\varepsilon} - \gamma + \ln 4\pi + 1 - \ln\frac{\lambda^2}{\mu^2}\right) \qquad (9.5.52)$$

Then the renormalized propagator is

$$G_{2ij}^{\mathrm{R}}(p) = \frac{\mathrm{i}\delta_{ij}}{\not p\left(1 - \frac{g_r^2}{(4\pi)^2}C_{\mathrm{F}}\ln\frac{-p^2}{\lambda^2}\right)} \qquad (9.5.53)$$

9.5.3.3. *The minimum subtraction scheme*

If only the singular term $1/\varepsilon$ is subtracted in the dimensional regularization, such subtraction scheme is called the minimum subtraction (MS) [27].

The least terms are subtracted in the minimum subtraction scheme, so it is widely used in perturbative calculations in QCD. In the minimum subtraction, the renormalization constants are taken as

$$Z_2 = 1 - \frac{g_r^2}{(4\pi)^2} C_{\mathrm{F}} \frac{1}{\varepsilon} \tag{9.5.54}$$

$$Z_2 Z_m = 1 - \frac{4g_r^2}{(4\pi)^2} C_{\mathrm{F}} \frac{1}{\varepsilon} \tag{9.5.55}$$

The renormalized propagator becomes

$$\begin{aligned} G_{2ij}^{\mathrm{R}}(p) = \mathrm{i}\delta_{ij} \Bigg\{ \not{p} \left[1 + \frac{g_r^2}{(4\pi)^2} C_{\mathrm{F}} \left(-\gamma + \ln 4\pi - 1 - 2\int_0^1 \mathrm{d}x (1-x) \ln \frac{L}{\mu^2} \right) \right] \\ -m_r \left[1 + \frac{4g_r^2}{(4\pi)^2} C_{\mathrm{F}} \left(-\gamma + \ln 4\pi - \frac{1}{2} - \int_0^1 \mathrm{d}x \ln \frac{L}{\mu^2} \right) \right] \Bigg\}^{-1} \end{aligned} \tag{9.5.56}$$

where $L = -x(1-x)p^2 + xm_r^2 - \mathrm{i}\varepsilon$.

9.5.3.4. *The modified minimum subtraction scheme*

From Eq. (9.5.40) we can see that, the divergent singular term $1/\varepsilon$ always emerges in association with Euler constant γ and the constant $\ln 4\pi$, which appears in the following combination

$$\frac{1}{\varepsilon} - \gamma + \ln 4\pi \tag{9.5.57}$$

Such combination always appears in the one-loop calculation of QCD, which stems from the dimensional regularization. Therefore, subtracting this combination as the whole divergent term in the renormalization procedure can make the calculation more concise. Such subtraction scheme is called the modified minimum subtraction, as ($\overline{\mathrm{MS}}$) [28]. In $\overline{\mathrm{MS}}$ scheme, the renormalization constants are

$$Z_2 = 1 - \frac{g_r^2}{(4\pi)^2} C_{\mathrm{F}} \left(\frac{1}{\varepsilon} - \gamma + \ln 4\pi \right) \tag{9.5.58}$$

$$Z_2 Z_m = 1 - \frac{4g_r^2}{(4\pi)^2} C_{\mathrm{F}} \left(\frac{1}{\varepsilon} - \gamma + \ln 4\pi \right) \tag{9.5.59}$$

and the renormalized quark propagator becomes

$$G^{\mathrm{R}}_{2ij}(p) = \mathrm{i}\delta_{ij}\left\{ \not{p}\left[1+\frac{g_r^2}{(4\pi)^2}C_{\mathrm{F}}\left(-1-2\int_0^1 \mathrm{d}x(1-x)\ln\frac{L}{\mu^2}\right)\right]\right.$$
$$\left. -m_r\left[1+\frac{4g_r^2}{(4\pi)^2}C_{\mathrm{F}}\left(-\frac{1}{2}-\int_0^1 \mathrm{d}x\ln\frac{L}{\mu^2}\right)\right]\right\}^{-1} \qquad (9.5.60)$$

where $L = -x(1-x)p^2 + xm_r^2 - \mathrm{i}\varepsilon$.

We have introduced four frequently-used subtraction schemes in the above, where we see that the renormalized propagators are not the same in different subtraction schemes, which shows the dependence of the perturbative calculation on the renormalization scheme. In fact, not only the propagators but also the Green functions, particle mass and coupling constants etc. are all dependent on the renormalization schemes.

9.5.4. *One-loop result of the renormalization constants*

The renormalization of quark one-loop self-energy part is discussed in Sec. 9.5.3 in Feynman gauge $\alpha = 1$, where the renormalization constants Z_2 and Z_m have been given. The results of all the renormalization constants up to one-loop order in arbitrary gauge will be given in this section, and the details of the calculation are omitted for simplicity.

9.5.4.1. *The gluon self-energy part*

The diagram for gluon self-energy is shown in Fig. 9.8.

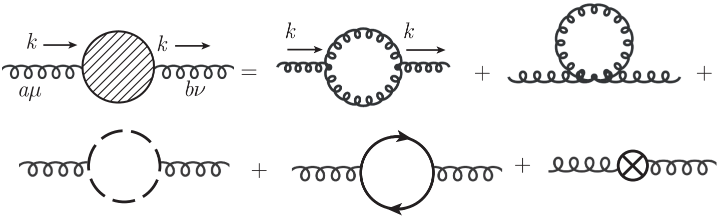


Fig. 9.8. One-loop diagram for gluon self-energy and the counter term.

The function for the gluon self-energy is

$$\prod_{\mu\nu}^{ab}(k) = \mathrm{i}\delta_{ab}(k_\mu k_\nu - k^2 g_{\mu\nu})\prod(k^2) \tag{9.5.61}$$

$$\prod(k^2) = \frac{g_r^2}{(4\pi)^2}\left[\frac{4}{3}T_{\mathrm{R}}N_{\mathrm{f}} - \frac{1}{2}C_{\mathrm{G}}\left(\frac{13}{3} - \alpha_r\right)\right]\frac{1}{\varepsilon} + Z_3 - 1 + \text{the finite terms} \tag{9.5.62}$$

where N_{f} is the number of quark flavors, T_{R} and C_{G} the constants defined in the following:

$$\begin{aligned}\mathrm{Tr}\left[T^a T^b\right] &= \delta_{ab}T_{\mathrm{R}} \\ f^{acd}f^{bcd} &= \delta_{ab}C_{\mathrm{G}}\end{aligned} \tag{9.5.63}$$

For SU(n) group, $T_{\mathrm{R}} = 1/2$, $C_{\mathrm{G}} = N$. The factor $(k_\mu k_\nu - k^2 g_{\mu\nu})$ in Eq. (9.5.61) is inevitable result of gauge invariance. It is exactly the appearance of this gauge invariant factor in $\prod_{\mu\nu}^{ab}(k)$ that causes the actual divergence degree to be 2 less than the superficial degree of divergence. This is an example of how symmetry can reduce the divergence of the momentum integral for a loop diagram. There is no mass term appearing in Eq. (9.5.61), so gluon is still massless after renormalization, which shows that the original massless particle cannot obtain mass from corrections of loop diagrams. It can be known from Eq. (9.5.62) that the renormalization Z_3 is

$$Z_3 = 1 - \frac{g_r^2}{(4\pi)^2}\left[\frac{4}{3}T_{\mathrm{R}}N_{\mathrm{f}} - \frac{1}{2}C_{\mathrm{G}}\left(\frac{13}{3} - \alpha_r\right)\right]\frac{1}{\varepsilon} + O(g_r^4) \tag{9.5.64}$$

in MS scheme.

9.5.4.2. *The self-energy of ghost field*

The ghost self-energy diagram is shown in Fig. 9.9.

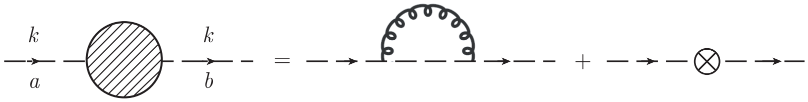


Fig. 9.9. The self-energy diagram of ghost and the counter term.

The self-energy function of ghost up to one-loop order in QCD is

$$\tilde{\prod}^{ab}(k) = \mathrm{i}\delta_{ab}k^2\left[-\frac{g_r^2}{(4\pi)^2}C_{\mathrm{G}}\frac{3-\alpha_r}{4}\frac{1}{\varepsilon}+\tilde{Z}_3-1\right] + \text{the finite terms} \tag{9.5.65}$$

So in MS scheme, the renormalization constant of the ghost field $\tilde{Z}_3$ is

$$\tilde{Z}_3 = 1+\frac{g_r^2}{(4\pi)^2}C_{\mathrm{G}}\frac{3-\alpha_r}{4}\frac{1}{\varepsilon}+O(g_r^4) \tag{9.5.66}$$

It can be seen from the result of the QCD one-loop correction for the ghost self-energy part that the result is proportional to k^2, and there is no mass term appearing in it. This indicates that it is impossible to generate mass for ghost particle through QCD higher-order corrections.

9.5.4.3. *The quark self-energy*

The quark self-energy diagram is shown in Fig. 9.10.

The quark self-energy function up to one-loop order in QCD is

$$\begin{aligned}\Sigma^{ij}(p) &= \mathrm{i}\delta_{ij}C_F\left[A\,\not{p}-Bm_r+(Z_2-1)\,\not{p}-(Z_2Z_m-1)m_r\right]\\ &\quad+\text{the finite terms}\\ A &= \frac{g_r^2}{(4\pi)^2}C_{\mathrm{F}}\alpha_r\frac{1}{\varepsilon}+O(g_r^4)\\ B &= \frac{g_r^2}{(4\pi)^2}C_{\mathrm{F}}(3+\alpha_r)\frac{1}{\varepsilon}+O(g_r^4)\end{aligned} \tag{9.5.67}$$

So in the MS scheme, the renormalization constants are

$$\begin{aligned}Z_2 &= 1+A+O(g_r^4)\\ Z_m &= 1+B-A+O(g_r^4)\end{aligned} \tag{9.5.68}$$

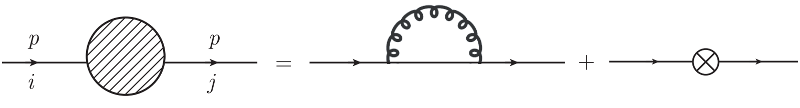


Fig. 9.10. The one-loop energy for quark self-energy and the counter term.

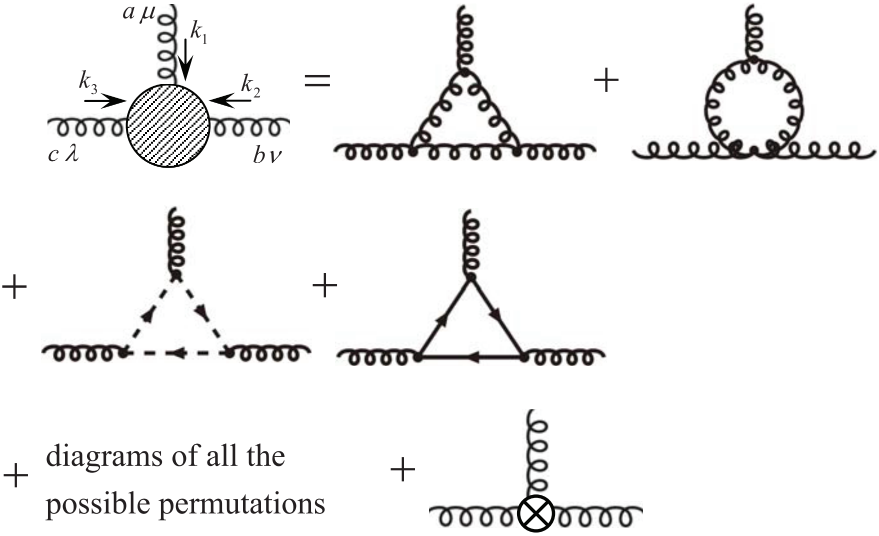


Fig. 9.11. The QCD one-loop correction diagram of three-gluon vertex and the counter term.

9.5.4.4. *The three-gluon vertex*

The one-loop diagram for three-gluon in QCD is depicted in Fig. 9.11. The QCD one-loop correction to three-gluon vertex is

$$\begin{aligned}\Lambda^{abc}_{\mu\nu\lambda}(k_1,k_2,k_3) = g_r f^{abc} V_{\mu\nu\lambda}(k_1,k_2,k_3)\\ \cdot\left[\frac{g_r^2}{(4\pi)^2}\left(C_{\mathrm{G}}\left(-\frac{17}{12}+\frac{3\alpha_r}{4}\right)+\frac{4}{3}T_{\mathrm{R}}N_{\mathrm{f}}\right)\frac{1}{\varepsilon}+Z_1-1\right]\\ +\text{the finite terms}\end{aligned} \tag{9.5.69}$$

From the above result, the renormalization constant Z_1 for three-gluon vertex in MS scheme can be obtained to be

$$Z_1 = 1-\frac{g_r^2}{(4\pi)^2}\left[C_{\mathrm{G}}\left(-\frac{17}{12}+\frac{3\alpha_r}{4}\right)+\frac{4}{3}T_{\mathrm{R}}N_{\mathrm{f}}\right]\frac{1}{\varepsilon}+O(g_r^4) \tag{9.5.70}$$

9.5.4.5. *The four-gluon vertex*

The QCD one-loop contribution to the four-gluon vertex is shown in Fig. 9.12.

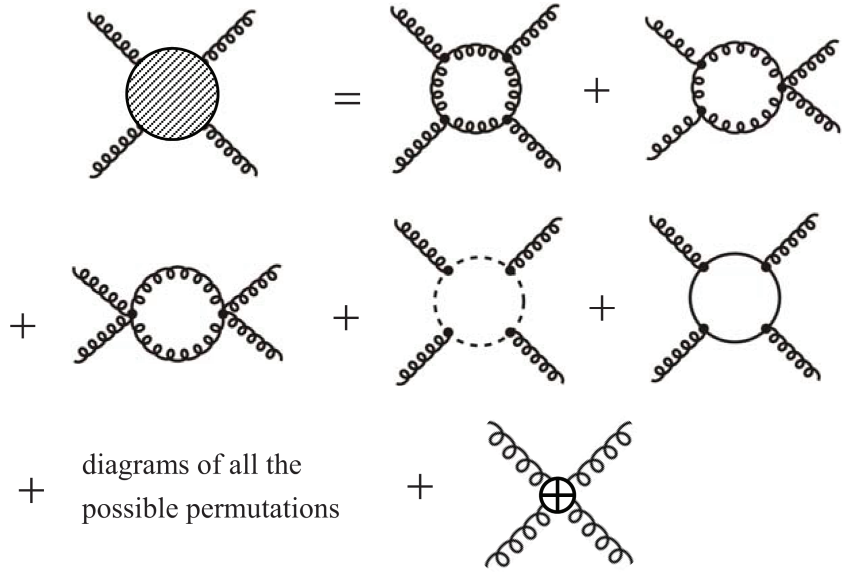


Fig. 9.12. The one-loop diagram in QCD for four-gluon vertex and the counter term.

The one-loop corrections to the four-gluon vertex is

$$\begin{aligned}\Lambda^{ijkl}_{\mu\nu\rho\sigma}(k_1,k_2,k_3,k_4) &= -\mathrm{i}g_r^2 W^{ijkl}_{\mu\nu\rho\sigma} \\ &\quad\cdot\left[\frac{g_r^2}{(4\pi)^2}\left(\left(-\frac{2}{3}+\alpha_r\right)C_{\mathrm{G}}+\frac{4}{3}T_{\mathrm{R}}N_{\mathrm{f}}\right)\frac{1}{\varepsilon}+Z_4-1\right] \\ &\quad+\text{the finite terms}\end{aligned} \tag{9.5.71}$$

from which one can obtain the renormalization constant for four-gluon vertex in MS scheme to be

$$Z_4 = 1-\frac{g_r^2}{(4\pi)^2}\left[\left(-\frac{2}{3}+\alpha_r\right)C_{\mathrm{G}}+\frac{4}{3}T_{\mathrm{R}}N_{\mathrm{f}}\right]\frac{1}{\varepsilon}+O(g_r^4) \tag{9.5.72}$$

9.5.4.6. *The ghost–gluon interaction vertex*

The one-loop diagram of gluon–ghost interaction vertex in QCD is depicted in Fig. 9.13.

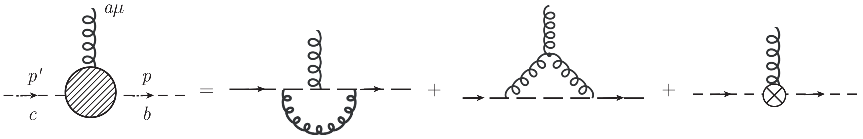


Fig. 9.13. The one-loop diagram of ghost–gluon interaction in QCD and the counter term.

The one-loop correction to the ghost–gluon interaction vertex in QCD is calculated to be

$$\tilde{\Lambda}_{\mu}^{abc}(k,p,p') = g_r f^{abc} p_\mu \left[\frac{g_r^2}{(4\pi)^2} C_{\mathrm{G}} \frac{\alpha_r}{2}\frac{1}{\varepsilon} + \tilde{Z}_1 - 1\right] + \text{the finite terms} \tag{9.5.73}$$

Therefore, in MS scheme, the renormalization constant for ghost–gluon interaction vertex can be obtained to be

$$\tilde{Z}_1 = 1 - \frac{g_r^2}{(4\pi)^2} C_{\mathrm{G}} \frac{\alpha_r}{2}\frac{1}{\varepsilon} + O(g_r^4) \tag{9.5.74}$$

9.5.4.7. *The quark–gluon vertex*

The one-loop correction diagram of quark–gluon interaction vertex in QCD can be shown in Fig. 9.14.

The QCD one-loop correction to quark–gluon interaction vertex is

$$\Lambda_{\mu}^{aij}(k,p,p') = \mathrm{i} g_r \gamma_\mu T_{ij}^a \left[\frac{g_r^2}{(4\pi)^2}\left(\frac{3+\alpha_r}{4} C_{\mathrm{G}} + \alpha_r C_{\mathrm{F}}\right)\frac{1}{\varepsilon} + Z_{1\mathrm{F}} - 1\right] + \text{the finite terms} \tag{9.5.75}$$

Then the renormalization constant $Z_{1\mathrm{F}}$ in MS scheme is

$$Z_{1\mathrm{F}} = 1 - \frac{g_r^2}{(4\pi)^2}\left(\frac{3+\alpha_r}{4} C_{\mathrm{G}} + \alpha_r C_{\mathrm{F}}\right)\frac{1}{\varepsilon} + O(g_r^4) \tag{9.5.76}$$

All the results for the seven superficially divergent Feynman diagrams in QCD are given above. From the divergent terms, eight renormalization constants, Z_3, $\tilde{Z}_3$, Z_2, Z_m, Z_1, Z_4, $\tilde{Z}_1$, $Z_{1\mathrm{F}}$ are determined. Seven among these eight constants are not independent in QCD. If the renormalization of QCD is self-consistent, these seven renormalization constants except Z_m should satisfy Slavnov–Taylor identity, i.e. Eq. (9.5.35). Using the renormalization constants calculated up to one-loop correction in QCD given

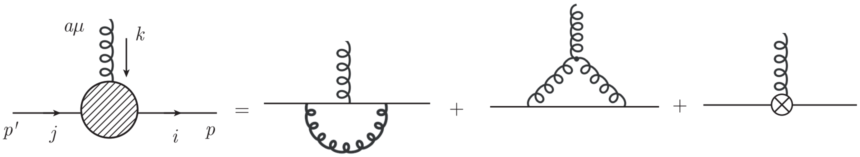


Fig. 9.14. The one-loop correction diagram for quark–gluon interaction and the counter term.

above, we can verify that they indeed satisfy

$$\frac{Z_1}{Z_3} = \frac{\tilde{Z}_1}{\tilde{Z}_3} = \frac{Z_{1\mathrm{F}}}{Z_2} = \frac{Z_4}{Z_1} = 1 - \frac{g_r^2}{(4\pi)^2} C_{\mathrm{G}} \frac{3+\alpha_r}{4} \frac{1}{\varepsilon} + O(g_r^4) \tag{9.5.77}$$

This is equivalent to say that the renormalizability of QCD is proved by actual calculation at one-loop level. Such a proof proceeds without taking into account the gauge symmetry of QCD. The $\mathrm{SU}(3)_C$ gauge symmetry disappears after gauge fixing terms added. However, if the transformation of the ghost fields are considered, gauge symmetry still hold for QCD after quantization. Such symmetry is called BRST symmetry. Identical equations between series of Green functions can be derived with BRST symmetry. These equations are referred to as generalized Ward–Takahashi identity, or Slavnov–Taylor identity. Equation (9.5.77) is a specific case of Slavnov–Taylor identity. Using gauge symmetry, renormalizability of QCD can be proved order by order, in principle, up to infinity orders. The reader can refer to Ref. [29] and the original literatures for the proof.

9.6. BRST Symmetry and the Generalized Ward–Takahashi Identity

9.6.1. *BRST symmetry*

The Lagrangian of QCD after quantization is given in Eq. (9.3.6), i.e.

$$\mathcal{L} = -\frac{1}{4}(F^a_{\mu\nu}F^{a\mu\nu}) + \bar{\psi}(\mathrm{i}\not{D} - m)\psi - \frac{1}{2\alpha}(\partial^\mu A^a_\mu)^2 + (\partial^\mu \bar{\chi}^a) D^{ab}_\mu \chi^b \tag{9.6.1}$$

Since the gauge fixing and Faddeev–Popov ghost terms are added in the Lagrangian, the Lagrangian is no longer invariant under gauge transformation

$$\psi' = U\psi, \quad U = \mathrm{e}^{-\mathrm{i}T^a\theta^a(x)} \tag{9.6.2}$$

$$T^a A'^a_\mu = UT^a A^a_\mu U^{-1} - \frac{\mathrm{i}}{g}(\partial_\mu U)U^{-1} \tag{9.6.3}$$

As $\theta^a(x)$ is infinitesimal parameter, Eqs. (9.6.2) and (9.6.3) become

$$\delta\psi = -\mathrm{i}T^a\theta^a\psi \tag{9.6.4}$$

$$\delta A_\mu^a = f^{abc}\theta^b A_\mu^c - \frac{1}{g}\partial_\mu\theta^a = -\frac{1}{g}D_\mu^{ab}\theta^b \tag{9.6.5}$$

where D_μ^{ab} is the covariant derivative in the adjoint representation, which is given in Eq. (9.3.3). Although the Lagrangian $\mathcal{L}$ as a whole is not gauge invariant, the first two terms are still invariant under the transformation of Eqs. (9.6.4) and (9.6.5). Can the Lagrangian $\mathcal{L}$ with the ghost and gauge fixing terms added be still gauge invariant if the ghost field is also appropriately changed under gauge transformation? In 1975–1976, Becchi, Rouet, Stora and Tyutin introduced a new symmetric transformation for the gauge fixed Lagrangian, which includes the transformation of the ghost fields. Under the new transformation, the quantized Lagrangian of QCD keeps invariant [30–32]. The basic idea of this transformation is to take the gauge transformation parameter in Eqs. (9.6.4) and (9.6.5) as

$$\theta^a(x) = -g\epsilon\chi^a(x) \tag{9.6.6}$$

where ϵ is an infinitesimal Grassmann number that is independent of the space–time coordinate

$$\{\epsilon, \chi^a(x)\} = 0 \tag{9.6.7}$$

It is known from Eq. (9.6.6) that ϵ must be restricted to be a number that can guarantee $\epsilon\chi^a$ to be real.

According to the way of selecting transformation parameter in Eq. (9.6.6), the local gauge transformation becomes

$$\delta\psi = \mathrm{i}\epsilon g T^a\chi^a\psi \tag{9.6.8}$$

$$\delta A_\mu^a = \epsilon D_\mu^{ab}\chi^b \tag{9.6.9}$$

Under the gauge transformation given above, the first two terms of $\mathcal{L}$ are certainly invariant. Next we need to determine how the ghost fields χ^a and $\bar{\chi}^a$ transform simultaneously as the quark and gauge fields transform according to Eqs. (9.6.8) and (9.6.9), then the added gauge fixing and ghost terms in the Lagrangian $\mathcal{L}$ in Eq. (9.6.1) can be guaranteed to be invariant.

Denote the infinitesimal transformation of the ghost fields as $\delta\chi^a$ and $\delta\bar{\chi}^a$, then the final two terms in Eq. (9.6.1) transform as

$$\delta\left(-\frac{1}{2\alpha}(\partial^\mu A_\mu^a)^2 + (\partial^\mu\bar{\chi}^a)D_\mu^{ab}\chi^b\right)$$
$$= -\frac{1}{\alpha}(\partial^\mu A_\mu^a)(\partial^\nu \delta A_\nu^a) + (\partial^\mu\delta\bar{\chi}^a)D_\mu^{ab}\chi^b + (\partial^\mu\bar{\chi}^a)\delta(D_\mu^{ab}\chi^b) \qquad (9.6.10)$$

and

$$\begin{aligned}\delta(D_\mu^{ab}\chi^b) &= \delta\left((\delta^{ab}\partial_\mu - gf^{abc}A_\mu^c)\chi^b\right)\\ &= \delta^{ab}\partial_\mu\delta\chi^b - gf^{abc}(\delta A_\mu^c)\chi^b - gf^{abc}A_\mu^c\delta\chi^b\\ &= \delta^{ab}\partial_\mu\delta\chi^b - \epsilon gf^{abc}(D_\mu^{cd}\chi^d)\chi^b - gf^{abc}A_\mu^c\delta\chi^b \qquad (9.6.11)\end{aligned}$$

Equation (9.6.9) has been used in the second to third steps in Eq. (9.6.11). Substitute Eqs. (9.6.9) and (9.6.11) into Eq. (9.6.10), and slightly simplify it, then after a 4-divergent term being dropped, the result of Eq. (9.6.10), i.e. the change of the whole Lagrangian $\mathcal{L}$ is

$$\begin{aligned}\delta\mathcal{L} = &-\left(\frac{\epsilon}{\alpha}(\partial^\mu A_\mu^a) + \delta\bar{\chi}^a\right)\partial^\nu(D_\nu^{ab}\chi^b)\\ &+ (\partial^\mu\bar{\chi}^a)\partial_\mu(\delta\chi^a + \frac{1}{2}\epsilon gf^{abc}\chi^b\chi^c)\\ &+ g(\partial^\mu\bar{\chi}^a)A_\mu^e(-f^{abe}\delta\chi^b - \epsilon gf^{abc}f^{cde}\chi^b\chi^d) \qquad (9.6.12)\end{aligned}$$

It is easily seen from the first two terms in Eq. (9.6.12) that, to keep $\delta\mathcal{L} = 0$, the change of the ghost fields should be

$$\delta\bar{\chi}^a = -\frac{\epsilon}{\alpha}(\partial^\mu A_\mu^a) \qquad (9.6.13)$$

$$\delta\chi^a = -\frac{1}{2}\epsilon gf^{abc}\chi^b\chi^c \qquad (9.6.14)$$

With the above two equations, the first two terms in Eq. (9.6.12) apparently vanish. We need to prove that the last term in Eq. (9.6.12) also vanish. Substitute Eq. (9.6.14) into the last term of Eq. (9.6.12), then the part in the parenthesis is

$$\begin{aligned}
&-f^{abe}\delta\chi^b-\epsilon g f^{abc}f^{cde}\chi^b\chi^d\\
&=\frac{1}{2}\epsilon g f^{abe}f^{bcd}\chi^c\chi^d-\epsilon g f^{abc}f^{cde}\chi^b\chi^d\\
&=\frac{1}{2}\epsilon g f^{ace}f^{cbd}\chi^b\chi^d-\epsilon g f^{abc}f^{cde}\chi^b\chi^d\\
&=\frac{1}{2}\epsilon g f^{ace}f^{cbd}\chi^b\chi^d-\frac{1}{2}\epsilon g(f^{abc}f^{cde}+f^{adc}f^{ceb})\chi^b\chi^d\\
&=-\frac{1}{2}\epsilon g(f^{aec}f^{cbd}+f^{abc}f^{cde}+f^{adc}f^{ceb})\chi^b\chi^d\\
&=0
\end{aligned}\tag{9.6.15}$$

Jacobi identity for the structure constants of the SU(3) group

$$f^{abc}f^{cde}+f^{bdc}f^{cae}+f^{dac}f^{cbe}=0\tag{9.6.16}$$

has been used in the last step of Eq. (9.6.15). Therefore, we have proved that, under the joint transformations of Eqs. (9.6.8), (9.6.9), (9.6.13) and (9.6.14), the quantized Lagrangian of QCD is invariant

$$\delta\mathcal{L}=0\tag{9.6.17}$$

Although the quantized Lagrangian of QCD $\mathcal{L}$ is no longer invariant under the classical local gauge transformation because of the existence of gauge fixing and ghost terms, it is still invariant under the new transformation of Eqs. (9.6.8), (9.6.9), (9.6.13) and (9.6.14). This is a new symmetry of QCD Lagrangian. It can be viewed as a new local gauge symmetry after quantization. It is called BRST symmetry, and the transformation in Eqs. (9.6.8), (9.6.9), (9.6.13) and (9.6.14) is referred to as BRST transformation, which can be summarized as follows:

$$\begin{aligned}
\delta\psi&=\mathrm{i}\epsilon g T^a\chi^a\psi,\quad \delta\bar{\psi}=-\mathrm{i}\epsilon g\chi^a\bar{\psi}T^a\\
\delta A^a_\mu&=\epsilon D^{ab}_\mu\chi^b\\
\delta\bar{\chi}^a&=-\frac{\epsilon}{\alpha}(\partial^\mu A^a_\mu)\\
\delta\chi^a&=-\frac{1}{2}\epsilon g f^{abc}\chi^b\chi^c
\end{aligned}\tag{9.6.18}$$

BRST symmetry is the basis of deriving the generalized Ward–Takahashi identity.

9.6.2. *The generalized Ward–Takahashi identity (Slavnov–Taylor identity)*

After the quantization of QCD Lagrangian, the classical gauge symmetry is replaced by BRST symmetry because of the existence of gauge fixing and ghost terms. Since the quantized Lagrangian still possesses a kind of symmetry, i.e. the BRST symmetry, some Green functions may not be completely independent of each other. That is, Green functions or generating functionals are imposed any restriction by BRST symmetry. Some of the identities existing between Green functions or generating functionals due to BRST symmetry are called the generalized Ward–Takahashi identities or Slavnov–Taylor identities. These identities play an important role in proving the renormalizability of QCD order by order. In history, 't Hooft first partially discussed the generalized Ward–Takahashi identity when he studied the renormalizability of the non-Abelian gauge field theory [16]. Slavnov and Taylor then did their own research thoroughly and independently [33, 34]. Now we introduce how to derive the generalized Ward–Takahashi identity from BRST symmetry.

Consider the generating functional of QCD given in Eq. (9.3.4)

$$\begin{aligned} Z\left[J,\xi,\bar{\xi},\eta,\bar{\eta}\right] = \int [\mathrm{d}A]\,[\mathrm{d}\chi]\,[\mathrm{d}\bar{\chi}]\,[\mathrm{d}\psi]\,\left[\mathrm{d}\bar{\psi}\right] \\ \times \exp\left\{\mathrm{i}\int \mathrm{d}x\left(\mathcal{L} + AJ + \bar{\chi}\xi + \bar{\xi}\chi + \bar{\psi}\eta + \bar{\eta}\psi\right)\right\} \end{aligned} \tag{9.6.19}$$

It is not difficult to prove that the functional integral measure in Eq. (9.6.19) is invariant under BRST transformation, i.e.

$$[\mathrm{d}A]\,[\mathrm{d}\chi]\,[\mathrm{d}\bar{\chi}]\,[\mathrm{d}\psi]\,\left[\mathrm{d}\bar{\psi}\right] = [\mathrm{d}A']\,[\mathrm{d}\chi']\,[\mathrm{d}\bar{\chi}']\,[\mathrm{d}\psi']\,\left[\mathrm{d}\bar{\psi}'\right] \tag{9.6.20}$$

where A', χ', $\bar{\chi}'$, ψ', $\bar{\psi}'$ are the relevant fields after BRST transformation. Here we shall not touch this proof further.

We have proved that the Lagrangian $\mathcal{L}$ is invariant under BRST transformation. So replace all the fields to the quantities after BRST transformation, i.e.

$$A \to A', \quad \chi \to \chi', \quad \bar{\chi} \to \bar{\chi}', \quad \psi \to \psi', \quad \bar{\psi} \to \bar{\psi}' \tag{9.6.21}$$

then the generating functional becomes

$$Z\left[J,\xi,\bar{\xi},\eta,\bar{\eta}\right] = \int [\mathrm{d}A]\,[\mathrm{d}\chi]\,[\mathrm{d}\bar{\chi}]\,[\mathrm{d}\psi]\,\left[\mathrm{d}\bar{\psi}\right] \times \exp\left\{\mathrm{i}\int \mathrm{d}x\,\left(\mathcal{L} + A'J + \bar{\chi}'\xi + \bar{\xi}\chi' + \bar{\psi}'\eta + \bar{\eta}\psi'\right)\right\} \tag{9.6.22}$$

under such replacement. The invariance of the functional integral measure and the Lagrangian $\mathcal{L}$ under BRST transformation has been considered in the above manipulation. Since applying BRST transformation to the generating functional is only a change of the symbol for the integral variables, the generated functional should, of course, remain unchanged, which means that Eqs. (9.6.22) and (9.6.19) should be equal to each other. This will lead to the Green functions derived from Eqs. (9.6.19) and (9.6.22) should be equal, i.e.

$$\langle 0|\,T[\hat{A}^a_\mu(x)\cdots]\,|0\rangle = \langle 0|\,T[\hat{A}'^a_\mu(x)\cdots]\,|0\rangle \tag{9.6.23}$$

Equation (9.6.23) can also be expressed as

$$\delta\,\langle 0|\,T[\hat{A}^a_\mu(x)\cdots]\,|0\rangle = 0 \tag{9.6.24}$$

Equation (9.6.24) means that Green function is invariant under BRST transformation. This is the generalized Ward–Takahashi identity for Green functions. The Slavnov–Taylor identity in the narrow sense given in Eq. (9.5.35) can be derived from the generalized Ward–Takahashi identity. This can be shown by selecting appropriate fields and considering the relevant Green functions. For example, the first equality of Eq. (9.5.35) can be proven by considering the identity held by the following Green function of gluon and ghost fields [33],

$$\delta\,\langle 0|\,T[\hat{A}^a_\mu(x)\hat{A}^b_\nu(y)\bar{\chi}^c(z)]\,|0\rangle = 0 \tag{9.6.25}$$

With Eq. (9.6.18), Eq. (9.6.25) can lead to

$$\begin{aligned}&\langle 0|\,T[(\hat{D}^{ad}_\mu\hat{\chi}^d(x))\hat{A}^b_\nu(y)\bar{\chi}^c(z)]\,|0\rangle\\ &\quad + \langle 0|\,T[\hat{A}^a_\mu(x)(\hat{D}^{bd}_\nu\hat{\chi}^d(y))\bar{\chi}^c(z)]\,|0\rangle\\ &\quad - \frac{1}{\alpha}\frac{\partial}{\partial z_\lambda}\,\langle 0|\,T[\hat{A}^a_\mu(x)\hat{A}^b_\nu(y)\hat{A}^c_\lambda(z)]\,|0\rangle = 0\end{aligned} \tag{9.6.26}$$

With the above equation, the first equality of Eq. (9.5.35) can be finally proven, while the other two equalities can be proven by selecting identities of Green functions for other appropriate fields.

The generalized Ward–Takahashi identity can also be expressed through restrictions to the generating functionals. Denote the relations between the fields before and after BRST transformation as

$$\begin{aligned} &A' = A + \delta A, \quad \bar{\chi}' = \bar{\chi} + \delta\bar{\chi}, \quad \chi' = \chi + \delta\chi, \\ &\psi' = \psi + \delta\psi, \quad \bar{\psi}' = \bar{\psi} + \delta\psi \end{aligned} \tag{9.6.27}$$

Subtracting Eq. (9.6.19) from Eq. (9.6.22) and preserving the result to the first order of the infinitesimal BRST transformation

$$\begin{aligned} &\int [\mathrm{d}A]\,[\mathrm{d}\chi]\,[\mathrm{d}\bar{\chi}]\,[\mathrm{d}\psi]\,\left[\mathrm{d}\bar{\psi}\right] \int \mathrm{d}^4x \left(\delta A J + \delta\bar{\chi}\xi + \bar{\xi}\delta\chi + \delta\bar{\psi}\eta + \bar{\eta}\delta\psi\right) \\ &\quad \times \exp\left\{\mathrm{i}\int \mathrm{d}^4x \left(\mathcal{L} + AJ + \bar{\chi}\xi + \bar{\xi}\chi + \bar{\psi}\eta + \bar{\eta}\psi\right)\right\} = 0 \end{aligned} \tag{9.6.28}$$

Substituting BRST transformation given in Eq. (9.6.18) into (9.6.28), and replacing each field by the relevant functional derivative, we can obtain the differential equation satisfied by the generating functional. But some of these differential operators in this equation are nonlinear, which makes this equation difficult to be applied. To get an linear differential equation, we need to introduce several new external source terms, which correspond to the compound fields $D_\mu^{ab}\chi^b$, $\frac{1}{2}gf^{abc}\chi^b\chi^c$ and $\mathrm{i}gT^a\chi^a\psi$. Replacing the external source terms in Eq. (9.6.19) with the following source term Σ [35]:

$$\begin{aligned} \Sigma \equiv{}& A_\mu^a J^{a\mu} + \bar{\chi}^a\xi^a + \bar{\xi}^a\chi^a + \bar{\psi}\eta + \bar{\eta}\psi \\ &+ K^{a\mu}D_\mu^{ab}\chi^b - \frac{1}{2}gf^{abc}\zeta^a\chi^b\chi^c \\ &+ \mathrm{i}gT^a\bar{\omega}\chi^a\psi - \mathrm{i}g\chi^a\bar{\psi}T^a\omega \end{aligned} \tag{9.6.29}$$

where $K^{a\mu}$, ζ^a, ω and $\bar{\omega}$ are new source terms, and $K^{a\mu}$ being Grassmann number, while ζ^a, ω and $\bar{\omega}$ are all ordinary numbers.

It should be noted that there is an important characteristic for BRST transformation, that is, the continuously twice BRST transformations on any fields results in zero, i.e.

$$\delta_{\mathrm{B}}(\delta_{\mathrm{B}}\phi) = 0 \tag{9.6.30}$$

where ϕ stands for any field among A_μ^a, χ^a, $\bar{\chi}^a$ and ψ, and the subscript of the variational symbol δ means BRST transformation. It is not difficult to prove this conclusion. We only need to continuously apply BRST transformation in Eq. (9.6.18) twice, and consider the fact that the structure constants of the SU(3) group obey Jacobi identity given in Eq. (9.6.16), then the proof is straightforward. The details of the proof is omitted here. Because of the characteristic given in Eq. (9.6.30), and considering that the newly introduced source terms in Σ actually correspond to once BRST transformation of A_μ^a, χ^a and ψ, $\bar{\psi}$, respectively, we can know that an extra BRST transformation of these new source terms in Σ should be zero because they are actually twice BRST transformations for the relevant fields, so

$$\delta\Sigma = \delta A_\mu^a J^{a\mu} + \delta\bar{\chi}^a \xi^a + \bar{\xi}^a \delta\chi^a + \delta\bar{\psi}\eta + \bar{\eta}\delta\psi \tag{9.6.31}$$

where we have set the convention that the source functions are BRST invariant.

If replacing the source terms in Eqs. (9.6.19) and (9.6.22) with Σ, repeating the above manipulations and using Eq. (9.6.31), then we can obtain another identity similar to Eq. (9.6.28)

$$\begin{aligned}&\int [\mathrm{d}A]\,[\mathrm{d}\chi]\,[\mathrm{d}\bar{\chi}]\,[\mathrm{d}\psi]\,\left[\mathrm{d}\bar{\psi}\right] \int \mathrm{d}^4x \left(\delta A J + \delta\bar{\chi}\xi + \bar{\xi}\delta\chi + \delta\bar{\psi}\eta + \bar{\eta}\delta\psi\right)\\ &\quad \times \exp\left\{\mathrm{i}\int \mathrm{d}^4x\,(\mathcal{L}+\Sigma)\right\} = 0\end{aligned} \tag{9.6.32}$$

Let us introduce a new symbol $\langle\ \rangle$, which is defined in the following. For any quantity F, there is

$$\langle F\rangle = \int [\mathrm{d}A]\,[\mathrm{d}\chi]\,[\mathrm{d}\bar{\chi}]\,[\mathrm{d}\psi]\,\left[\mathrm{d}\bar{\psi}\right] F \exp\left\{\mathrm{i}\int \mathrm{d}^4x\,(\mathcal{L}+\Sigma)\right\} \tag{9.6.33}$$

then Eq. (9.6.32) can be simply written as

$$\int \mathrm{d}^4x \left(\langle\delta A\rangle J + \langle\delta\bar{\chi}\rangle \xi + \bar{\xi}\langle\delta\chi\rangle + \left\langle\delta\bar{\psi}\right\rangle \eta + \bar{\eta}\langle\delta\psi\rangle\right) = 0 \tag{9.6.34}$$

In the case that the source terms in Eq. (9.6.19) are replaced with Σ, we have

$$\begin{aligned}&\epsilon\frac{\delta Z}{\delta K_\mu^a} = \mathrm{i}\,\langle\delta A^{a\mu}\rangle\,, \quad \epsilon\frac{\delta Z}{\delta\zeta^a} = \mathrm{i}\,\langle\delta\chi^a\rangle\\ &\epsilon\frac{\delta Z}{\delta\omega} = \mathrm{i}\left\langle\delta\bar{\psi}\right\rangle, \quad \epsilon\frac{\delta Z}{\delta\bar{\omega}} = \mathrm{i}\,\langle\delta\psi\rangle\end{aligned} \tag{9.6.35}$$

With the above equations, Eq. (9.6.34) can be written as

$$\int \mathrm{d}^4x \left(\frac{\delta Z}{\delta K^a_\mu}J^a_\mu - \frac{1}{\alpha}\left(\partial_\mu \frac{\delta Z}{\delta J^a_\mu}\xi^a\right) + \bar{\xi}^a\frac{\delta Z}{\delta \zeta^a} + \frac{\delta Z}{\delta \omega}\eta + \bar{\eta}\frac{\delta Z}{\delta \bar{\omega}}\right) = 0 \quad (9.6.36)$$

Equation (9.6.36) is just the generalized Ward–takahashi identity expressed with generating functional Z.

We have given the generating functional for the connected Green function W in Eq. (9.3.36), which is related to Z by $Z = \exp(\mathrm{i}W)$. Substituting this relation into Eq. (9.6.36), we can obtain the identity satisfied by W, which is similar to Eq. (9.6.36)

$$\int \mathrm{d}^4x \left(\frac{\delta W}{\delta K^a_\mu}J^a_\mu - \frac{1}{\alpha}\left(\partial_\mu \frac{\delta W}{\delta J^a_\mu}\xi^a\right) + \bar{\xi}^a\frac{\delta W}{\delta \zeta^a} + \frac{\delta W}{\delta \omega}\eta + \bar{\eta}\frac{\delta W}{\delta \bar{\omega}}\right) = 0 \quad (9.6.37)$$

Equation (9.6.37) is the generalized Ward–takahashi identity satisfied by the connected Green function W.

The generalized Ward–takahashi identity can be also expressed using equations satisfied by the effective action, which is more suitable for studying the renormalizability of gauge theory. The effective action is the generating functional of one-particle irreducible Green function, which is defined through the generating functional W

$$\begin{aligned}&\Gamma\left[\upsilon_A, \upsilon_\chi, \upsilon_{\bar{\chi}}, \upsilon_\psi, \upsilon_{\bar{\psi}}; K, \zeta, \omega, \bar{\omega}\right]\\ &\quad = W\left[J, \xi, \bar{\xi}, \eta, \bar{\eta}, K, \zeta, \omega, \bar{\omega}\right]\\ &\qquad - \int \mathrm{d}^4x \left(\upsilon_A^{a\mu}J^a_\mu + \upsilon^a_{\bar{\chi}}\xi^a + \bar{\xi}^a\upsilon^a_\chi + \upsilon_{\bar{\psi}}\eta + \bar{\eta}\upsilon_\psi\right)\end{aligned} \quad (9.6.38)$$

where the new variables υ_A, υ_χ, $\upsilon_{\bar{\chi}}$, υ_ψ and $\upsilon_{\bar{\psi}}$ are defined as

$$\upsilon_A^{a\mu} = \frac{\delta W}{\delta J^a_\mu}, \quad \upsilon^a_\chi = \frac{\delta W}{\delta \bar{\xi}^a}, \quad \upsilon^a_{\bar{\chi}} = -\frac{\delta W}{\delta \xi^a}, \quad \upsilon_{\bar{\psi}} = -\frac{\delta W}{\delta \eta}, \quad \upsilon_\psi = \frac{\delta W}{\delta \eta} \quad (9.6.39)$$

According to the definition of Γ, i.e. Eq. (9.6.38), we can transform the identity (9.6.37) for W to the identity for Γ

$$\int \mathrm{d}^4x \left[\frac{\delta \Gamma}{\delta K^a_\mu}\frac{\delta \Gamma}{\delta \upsilon_A^{a\mu}} - \frac{1}{\alpha}(\partial_\mu \upsilon_A^{a\mu})\frac{\delta \Gamma}{\delta \upsilon^a_{\bar{\chi}}} - \frac{\delta \Gamma}{\delta \upsilon^a_\chi}\frac{\delta \Gamma}{\delta \zeta^a} + \frac{\delta \Gamma}{\delta \omega}\frac{\delta \Gamma}{\delta \upsilon_{\bar{\psi}}} + \frac{\delta \Gamma}{\delta \upsilon_\psi}\frac{\delta \Gamma}{\delta \bar{\omega}}\right] = 0 \quad (9.6.40)$$

Considering the effective Lagrangian $\mathcal{L}+\Sigma$, one can obtain the equation of motion for χ^a

$$\partial^\mu (D_\mu^{ab}\chi^b) - \xi^a = 0 \tag{9.6.41}$$

Then we have

$$\left\langle \left[\partial^\mu (D_\mu^{ab}\chi^b) - \xi^a\right] \right\rangle = 0 \tag{9.6.42}$$

Since ξ^a in the second term of Eq. (9.6.42) $\langle \xi^a \rangle$ is independent of the functional integral, it can be taken out of the integral calculation. Therefore

$$\langle \xi^a \rangle = \xi^a \langle 1 \rangle = \xi^a Z \tag{9.6.43}$$

With the result of Eq. (9.6.43), Eq. (9.6.42) can be written as

$$\partial_\mu \frac{\delta Z}{\mathrm{i}\delta K_\mu^a} = \xi^a Z \tag{9.6.44}$$

Considering the relation between Z, W and Γ, we can finally write Eq. (9.6.44) as

$$\partial_\mu \frac{\delta \Gamma}{\delta K_\mu^a} = -\frac{\delta \Gamma}{\delta \upsilon_{\bar{\chi}}^a} \tag{9.6.45}$$

Substituting Eq. (9.6.45) into the second term of Eq. (9.6.40), and doing integration by part, we can obtain

$$\int \mathrm{d}^4 x \left[\frac{\delta \Gamma}{\delta K_\mu^a} \left(\frac{\delta \Gamma}{\delta \upsilon_A^{a\mu}} - \frac{1}{\alpha} \partial_\mu \partial_\nu \upsilon_A^{a\nu} \right) - \frac{\delta \Gamma}{\delta \upsilon_\chi^a} \frac{\delta \Gamma}{\delta \zeta^a} + \frac{\delta \Gamma}{\delta \omega} \frac{\delta \Gamma}{\delta \upsilon_{\bar{\psi}}} + \frac{\delta \Gamma}{\delta \upsilon_\psi} \frac{\delta \Gamma}{\delta \bar{\omega}} \right] = 0 \tag{9.6.46}$$

Equation (9.6.46) can be simplified by defining a new generating functional $\tilde{\Gamma}$, with

$$\tilde{\Gamma} = \Gamma + \frac{1}{2\alpha} \int \mathrm{d}^4 x (\partial_\nu \upsilon_A^{a\nu})^2 \tag{9.6.47}$$

With the new generating functional $\tilde{\Gamma}$, the generalized Ward–Takahashi identity can be written as

$$\int \mathrm{d}^4 x \left[\frac{\delta \tilde{\Gamma}}{\delta K_\mu^a} \frac{\delta \tilde{\Gamma}}{\delta \upsilon_A^{a\mu}} - \frac{\delta \tilde{\Gamma}}{\delta \upsilon_\chi^a} \frac{\delta \tilde{\Gamma}}{\delta \zeta^a} + \frac{\delta \tilde{\Gamma}}{\delta \omega} \frac{\delta \tilde{\Gamma}}{\delta \upsilon_{\bar{\psi}}} + \frac{\delta \tilde{\Gamma}}{\delta \upsilon_\psi} \frac{\delta \tilde{\Gamma}}{\delta \bar{\omega}} \right] = 0 \tag{9.6.48}$$

Equation (9.6.48) has been frequently used in proving the renormalizability of gauge theories.

9.7. The Renormalization Group Equation and the Solution

9.7.1. *The renormalization group*

We have known from the renormalization procedure that the divergences appearing in the calculation of high-order contributions to Green function are subtracted step by step. This will generate the uncertainty of how to define the term of infinity, namely how many finite terms are eliminated when subtracting the infinite pole at the same time. This is what we discussed earlier, the uncertainty resulted from renormalization scheme dependence.

Another uncertainty in renormalization is the inevitable occurrence of an energy scale μ, called renormalization scale, in the process of subtraction of divergence. Some different subtraction schemes correspond to the choice of renormalization scale. For example, the subtraction condition in the on-shell subtraction scheme is chosen on the mass-shell of the particles, which corresponds to set the condition for fixing the renormalization scale μ on the mass-shell. Meanwhile, the condition for fixing the renormalization scale is set on any off-shell value for the four-momentum squared of the particles in off-shell renormalization scheme, i.e. $p^2 = -\mu^2 \neq m^2$. In the minimum subtraction scheme (MS) or the modified minimum subtraction scheme ($\overline{\text{MS}}$), there naturally appears an energy scale μ, which is the mass dimension of the coupling constant in D-dimensional space–time, $g = g_0\mu^{\varepsilon}$. Here μ is the renormalization scale in MS or $\overline{\text{MS}}$ subtraction scheme. Except the on-shell renormalization, the scale μ in other renormalization schemes is completely an undetermined arbitrary parameter, which is a residue after subtracting the infinity in the Green functions.

Since it is inevitable for such kind of uncertainty to appear in renormalization procedure, there may be different forms of expression corresponding to different subtraction schemes and different scale selections for a physical quantity. Then a natural question is if these different forms of expression are equivalent. These different expressions for a single quantity must be equivalent because they are derived from the same Lagrangian. This requires the physical quantities calculated in quantum field theory to be invariant for different subtraction schemes and different renormalization scales.

In different subtraction schemes and on different renormalization scales, the different expressions of a physical quantity must differ by a finite transformation, which is called finite renormalization. From Eqs. (9.5.29) and

(9.5.30), the relation between the renormalized and unrenormalized bare quantities of fields, coupling constant, quark mass and gauge fixing parameter are

$$\psi = Z_2^{1/2}\psi_{\rm r}, \quad A_\mu^a = Z_3^{1/2}A_{{\rm r}\mu}^a, \quad \chi^a = \tilde{Z}_3^{1/2}\chi_{\rm r}^a, \quad \bar{\chi}^a = \tilde{Z}_3^{1/2}\bar{\chi}_{\rm r}^a \tag{9.7.1}$$

$$g = Z_g g_{\rm r}, \quad m = Z_m m_{\rm r}, \quad \alpha = Z_3\alpha_{\rm r} \tag{9.7.2}$$

Corresponding to the renormalization of fields, the source functions in the generating functional, Eq. (9.3.4), need also to be renormalized

$$\begin{gathered} J_\mu^a = Z_3^{-1/2}J_{{\rm r}\mu}^a, \quad \xi = \tilde{Z}_3^{-1/2}\xi_{\rm r}, \quad \bar{\xi} = \tilde{Z}_3^{-1/2}\bar{\xi}_{\rm r} \\ \eta = Z_2^{-1/2}\eta_{\rm r}, \quad \bar{\eta} = Z_2^{-1/2}\bar{\eta}_{\rm r} \end{gathered} \tag{9.7.3}$$

The subscript "r" for each source function in the above stands for renormalization. The renormalized quantities in Eq. (9.7.3) are selected in this way so that the source terms in the generating functional are unchanged after renormalization. In order to make the discussion simpler, in the case of no ambiguity, we write Eq. (9.7.3) as follows:

$$J_i = Z_i^{-1/2}J_{i{\rm r}} \tag{9.7.4}$$

where the subscript "i" stands for all the possible source functions. We can select any kind of subtraction scheme and arbitrary renormalization scale because of the arbitrariness of renormalization procedure. One can also select another renormalization scheme different from the scheme mentioned above. In the new renormalization scheme, the relation between the bare and renormalized quantities can be denoted as

$$\psi = Z'^{1/2}_2\psi'_{\rm r}, \quad A_\mu^a = Z'^{1/2}_3 A'^a_{{\rm r}\mu}, \quad \chi^a = \tilde{Z}'^{1/2}_3\chi'^a_{\rm r}, \quad \bar{\chi}^a = \tilde{Z}'^{1/2}_3\bar{\chi}'^a_{\rm r} \tag{9.7.5}$$

$$g = Z'_g g'_{\rm r}, \quad m = Z'_m m'_{\rm r}, \quad \alpha = Z'_3\alpha'_{\rm r} \tag{9.7.6}$$

$$J_i = Z'^{-1/2}_i J'_{i{\rm r}} \tag{9.7.7}$$

where all the subtraction scheme and renormalization scale can be changed. Suppose that the renormalization scale is changed from μ in Eqs. (9.7.1)–(9.7.4) to μ' in the new renormalization scheme. It can be known by comparing these two different renormalization processes that the renormalized quantities in these two different renormalization schemes differ by a transformation

$$\psi'_{\rm r} = z_2^{1/2}\psi_{\rm r}, \quad A'^{a}_{{\rm r}\mu} = z_3^{1/2}A^{a}_{{\rm r}\mu}, \quad \chi'^{a}_{\rm r} = \tilde{z}_3^{1/2}\chi^{a}_{\rm r},$$
$$\bar{\chi}'^{a}_{\rm r} = \tilde{z}_3^{1/2}\bar{\chi}^{a}_{\rm r} \tag{9.7.8}$$

$$g'_{\rm r} = z_g g_{\rm r}, \quad m'_{\rm r} = z_m m_{\rm r}, \quad \alpha'_{\rm r} = z_3\alpha_{\rm r} \tag{9.7.9}$$

$$J'_{i{\rm r}} = z_i^{-1/2}J_{i{\rm r}} \tag{9.7.10}$$

where the definitions of z_2, z_3, $\tilde{z}_3$, z_g, z_m are

$$\begin{aligned} z_2 &= Z_2/Z'_2, \quad z_3 = Z_3/Z'_3, \quad \tilde{z}_3 = \tilde{Z}_3/\tilde{Z}'_3 \\ z_g &= Z_g/Z'_g, \quad z_m = Z_m/Z'_m, \quad z_i = Z_i/Z'_i \end{aligned} \tag{9.7.11}$$

The constants for the transformation in Eq. (9.7.11) are all finite, because the infinities contained in the renormalization constants in the numerator and denominators can cancel. This is why the transformation between the two renormalization schemes is called finite renormalization.

Because the predicted values of the same physical quantity must be the same in two different renormalization schemes, the generating functional must remain the same under different renormalization schemes

$$Z'_{\rm r}\left[J'_{i{\rm r}}, g'_{\rm r}, m'_{\rm r}, \mu'\right] = Z_{\rm r}\left[J_{i{\rm r}}, g_{\rm r}, m_{\rm r}, \mu\right] \tag{9.7.12}$$

and the generating functional for the connected Green function must also be equal

$$W'_{\rm r}\left[J'_{i{\rm r}}, g'_{\rm r}, m'_{\rm r}, \mu'\right] = W_{\rm r}\left[J_{i{\rm r}}, g_{\rm r}, m_{\rm r}, \mu\right] \tag{9.7.13}$$

Green function can be obtained by taking the functional derivative of the generating functional with respect to the source function $J_{i{\rm r}}$, and the S matrix element describing particle scattering can be obtained from Green function: $S\left[p, g_{\rm r}, m_{\rm r}, \mu\right]$, where p represents the set of momenta of all the particles involved in the scattering. Because of the uniqueness for the prediction of physical quantity, it is required that the S matrix element is invariant under the transformation of the finite renormalization

$$S'\left[p, g'_{\rm r}, m'_{\rm r}, \mu'\right] = S\left[p, g_{\rm r}, m_{\rm r}, \mu\right] \tag{9.7.14}$$

S and S' represent the matrix element S calculated in two different renormalization schemes, respectively. They are functions of the renormalized coupling constant and quark mass. Certainly, the form of the function of S' depending on $g'_{\rm r}$ and $m'_{\rm r}$ can be apparently different from that of

$S\left[p, g_{\mathrm{r}}, m_{\mathrm{r}}, \mu\right]$. But the values of them must be the same. In general, we calculate the scattering cross section of particles through S matrix element $S\left[p, g_{\mathrm{r}}, m_{\mathrm{r}}, \mu\right]$ and obtain the values of g_{r} and m_{r} by comparing the calculated cross section with experimental result, then we can predict the variation of the scattering cross section with the change of the momentum p. We can also calculate the scattering cross section through the S matrix element $S'\left[p, g'_{\mathrm{r}}, m'_{\mathrm{r}}, \mu'\right]$ in another renormalization scheme, and obtain g'_{r} and m'_{r} that are consistent with experiment by comparing the calculation and experimental result, and then predict the dependence of the cross section on the momentum p. It is worth pointing out that the scattering cross sections of these two predictions must give the same dependence on momentum p. g'_{r}, m'_{r} and g_{r}, m_{r} must be consistent with the relation of finite renormalization, i.e. Eq. (9.7.9). If the S matrix element $S\left[p, g_{\mathrm{r}}, m_{\mathrm{r}}, \mu\right]$ can be calculated exactly, then the above statement can be verified. But we can only calculate the S matrix elements of the scattering processes order by order in perturbation theory. To guarantee the perturbative expansion effective, we always take a renormalization scale in which the fitted coupling constant is small enough. Since the higher-order contribution is dropped in perturbative expansion method, $S'\left[p, g'_{\mathrm{r}}, m'_{\mathrm{r}}, \mu'\right]$ and $S\left[p, g_{\mathrm{r}}, m_{\mathrm{r}}, \mu\right]$ are not exactly equal to each other. They are only approximately equal within the accuracy of calculation. When we calculate perturbatively to the order n, $S'\left[p, g'_{\mathrm{r}}, m'_{\mathrm{r}}, \mu'\right]$ and $S\left[p, g_{\mathrm{r}}, m_{\mathrm{r}}, \mu\right]$ may differ by a bit, but the difference is at the order higher than g_{r}^{n+2}, i.e.

$$S'\left[p, g'_{\mathrm{r}}, m'_{\mathrm{r}}, \mu'\right] - S\left[p, g_{\mathrm{r}}, m_{\mathrm{r}}, \mu\right] = O(g_{\mathrm{r}}^{n+2}) \tag{9.7.15}$$

This difference will result in a dependency on the renormalization scheme for the results of the perturbation calculation. In theory the perturbation expansion is always calculated to higher orders so that the difference can be reduced order by order.

Next we shall continue to discuss the topics about renormalization.

9.7.1.1. *The finite renormalization of the coupling constant*

The renormalized coupling constant g_{r} depends on the renormalization scale, because the renormalization constant Z_g is renormalization scale dependent. The relation between g_{r} and the bare coupling constant g is

$$g_{\mathrm{r}}(\mu) = Z_g^{-1}(\mu) g \tag{9.7.16}$$

where we explicitly write the dependence of g_{r} and Z_g on the scale μ. In another renormalization scheme, we use μ' to tag this new renormalization procedure. The new renormalized coupling constant is

$$g_{\mathrm{r}}(\mu') = Z_g^{-1}(\mu')g \tag{9.7.17}$$

Comparing Eqs. (9.7.17) and (9.7.16), we can obtain the transformation from $g_{\mathrm{r}}(\mu) \to g_{\mathrm{r}}(\mu')$ as

$$g_{\mathrm{r}}(\mu') = z_g(\mu', \mu)g_{\mathrm{r}}(\mu) \tag{9.7.18}$$

$$z_g(\mu', \mu) = Z_g(\mu)/Z_g(\mu') \tag{9.7.19}$$

$Z_g(\mu', \mu)$ in Eq. (9.7.18) results in a finite renormalization from $g_{\mathrm{r}}(\mu) \to g_{\mathrm{r}}(\mu')$. All $z_g(\mu', \mu)$ obeying Eqs. (9.7.18) and (9.7.19) constitute a set, which is denoted as $\{z_g(\mu', \mu)\}$. This set satisfies all the characteristic of group [36].

Firstly, one can define the product of two elements $z_g(\mu'', \mu')$ and $z_g(\mu', \mu)$ in this set as

$$z_g(\mu'', \mu')z_g(\mu', \mu) \tag{9.7.20}$$

It should be noted that, when two finite renormalization elements $z_g(\mu'', \mu')$ and $z_g(\mu', \mu)$ are multiplied, there must be a common scale μ' within them. This product stands for the continuous transformation from the scale $\mu \to \mu' \to \mu''$. It can be seen from Eq. (9.7.19) that this product is just the transformation of $z_g(\mu'', \mu')$, which is the scale transformation from $\mu \to \mu''$. Therefore, this product also belongs to the set $\{z_g(\mu', \mu)\}$.

Secondly, one can define the inverse element of $z_g(\mu', \mu)$ as

$$z_g^{-1}(\mu', \mu) = z_g(\mu, \mu') \tag{9.7.21}$$

which stands for the transformation of the scale from $\mu' \to \mu$. In addition, it can be apparently seen from Eq. (9.7.19) that

$$z_g(\mu, \mu) = 1 \tag{9.7.22}$$

which is the identity of the set of finite renormalization. Therefore, the set conforms with all the characteristics of a group. It constitutes a group, and it is known from the product of the group elements given in Eq. (9.7.20) that the group elements are commutative. Therefore, this group is Abelian.

9.7.1.2. *The finite renormalization of quark mass*

The above discussion on the renormalization of the coupling constant can also applies to the quark mass renormalization. The renormalized quark mass m_r also depends on the renormalization scale

$$m_\mathrm{r}(\mu) = Z_m^{-1}(\mu)m \tag{9.7.23}$$

In another renormalization scheme

$$m_\mathrm{r}(\mu') = Z_m^{-1}(\mu')m \tag{9.7.24}$$

so we have

$$m_\mathrm{r}(\mu') = z_m(\mu', \mu)m_\mathrm{r}(\mu) \tag{9.7.25}$$

$$z_m(\mu', \mu) = Z_m(\mu)/Z_m(\mu') \tag{9.7.26}$$

The set of finite renormalization of quark mass $\{z_m(\mu', \mu)\}$ also constitutes an Abelian group.

9.7.1.3. *The finite renormalization of green function*

We take n gluon fields $A^{a_1}_{\mu_1}(x_1), A^{a_2}_{\mu_2}(x_2), \ldots, A^{a_n}_{\mu_n}(x_n)$ as an example to show the finite renormalization of Green function. The Green function of n gluon field operators is

$$G_n(x_1, \ldots, x_n) = \langle 0|\, T[\hat{A}^{a_1}_{\mu_1}(x_1)\hat{A}^{a_2}_{\mu_2}(x_2)\cdots\hat{A}^{a_n}_{\mu_n}(x_n)]\,|0\rangle \tag{9.7.27}$$

Here the Lorentz indices in G_n have been omitted, since they are not the focus of the discussion for the moment. The n-point Green function in the momentum space is denoted by $\tilde{G}_n$, then we have

$$\begin{aligned}&(2\pi)^4\delta^4(p_1 + \cdots + p_n)\tilde{G}_n(p, g, m, \alpha)\\ &\quad = \int \mathrm{d}^4x_1 \cdots \mathrm{d}^4x_n \mathrm{e}^{-\mathrm{i}(p_1\cdot x_1 + \cdots + p_n\cdot x_n)}G_n(x_1, \ldots, x_n)\end{aligned} \tag{9.7.28}$$

where p represents the whole set of momenta $\{p_1, p_1, \ldots, p_n\}$, g is the coupling constant, m the quark mass, and α the gauge fixing parameter. Although the Green function we discuss is for n gluon field operators, the Green function of gluon fields is also the function of quark mass, because quark field can contribute to the Green function of gluon fields as internal line.

The connected Green function is the result of removing the contribution of unconnected Feynman diagrams from the Green function, which is denoted by a superscript "c"

$$G_n^{\mathrm{c}}(x_1,\ldots,x_n)=\langle 0|\,T[\hat{A}_{\mu_1}^{a_1}(x_1)\hat{A}_{\mu_2}^{a_2}(x_2)\cdots\hat{A}_{\mu_n}^{a_n}(x_n)]\,|0\rangle_{\mathrm{c}} \tag{9.7.29}$$

The connected Green function can be obtained by taking the functional derivative of the connected generating functional $W\,[J,g,m,\alpha]$

$$G_n^{\mathrm{c}}(x_1,\ldots,x_n)=(-\mathrm{i})^{n-1}\left.\frac{\delta^n W\,[J,g,m,\alpha]}{\delta J_{\mu_1}^{a_1}(x_1)\cdots\delta J_{\mu_n}^{a_n}(x_n)}\right|_{J=0} \tag{9.7.30}$$

It can be seen from the renormalization of gluon field $A_\mu^a=Z_3^{1/2}A_{\mathrm{r}\mu}^a$ that the relations between the connected Green function after renormalization and the unrenormalized Green function are

$$G_{\mathrm{r}n}^{\mathrm{c}}=Z_3^{-n/2}G_n^{\mathrm{c}} \tag{9.7.31}$$

$$\tilde{G}_{\mathrm{r}n}^{\mathrm{c}}=Z_3^{-n/2}\tilde{G}_n^{\mathrm{c}} \tag{9.7.32}$$

The truncated connected Green function $\tilde{G}_n^{\mathrm{tc}}$ corresponds to the Feynman amplitude of particle scattering process, and $\tilde{G}_n^{\mathrm{tc}}$ is defined by the following equation:

$$\begin{aligned}&(2\pi)^4\delta^4(p_1+\cdots+p_n)\tilde{G}_2(p_1)\cdots\tilde{G}_2(p_n)\tilde{G}_n^{\mathrm{tc}}(p,g,m,\alpha)\\&\quad=\int\mathrm{d}^4x_1\cdots\mathrm{d}^4x_n\mathrm{e}^{-\mathrm{i}(p_1\cdot x_1+\cdots+p_n\cdot x_n)}G_n^{\mathrm{c}}(x_1+\cdots+x_n)\end{aligned} \tag{9.7.33}$$

Using Eqs. (9.7.31) and (9.7.32) in Eq. (9.7.33), and cancel the excess $Z_3^{1/2}$ on both sides, we can obtain

$$\begin{aligned}&(2\pi)^4\delta^4(p_1+\ldots+p_n)\tilde{G}_{\mathrm{r}2}(p_1)\cdots\tilde{G}_{\mathrm{r}2}(p_n)Z_3^{n/2}\tilde{G}_n^{\mathrm{tc}}(p,g,m,\alpha)\\&\quad=\int\mathrm{d}^4x_1\cdots\mathrm{d}^4x_n\mathrm{e}^{-\mathrm{i}(p_1\cdot x_1+\cdots+p_n\cdot x_n)}G_{\mathrm{r}n}^{\mathrm{c}}(x_1,\ldots,x_n)\end{aligned} \tag{9.7.34}$$

while the renormalized truncated and connected Green function $\tilde{G}_{\mathrm{r}n}^{\mathrm{tc}}$ is defined through the following equation:

$$\begin{aligned}&(2\pi)^4\delta^4(p_1+\cdots+p_n)\tilde{G}_{\mathrm{r}2}(p_1)\cdots\tilde{G}_{\mathrm{r}2}(p_n)\tilde{G}_{\mathrm{r}n}^{\mathrm{tc}}(p,g_{\mathrm{r}},m_{\mathrm{r}},\alpha_{\mathrm{r}})\\&\quad=\int\mathrm{d}^4x_1\cdots\mathrm{d}^4x_n\mathrm{e}^{-\mathrm{i}(p_1\cdot x_1+\cdots+p_n\cdot x_n)}G_{\mathrm{r}n}^{\mathrm{c}}(x_1,\ldots,x_n)\end{aligned} \tag{9.7.35}$$

Comparing Eqs. (9.7.34) and (9.7.35) results in

$$\tilde{G}^{\mathrm{tc}}_{\mathrm{r}n}(p,g_{\mathrm{r}},m_{\mathrm{r}},\alpha_{\mathrm{r}}) = Z_3^{n/2}(\mu)\tilde{G}^{\mathrm{tc}}_n(p,g,m,\alpha) \tag{9.7.36}$$

The renormalized truncated and connected Green function is defined as the renormalized Feynman amplitude F_n. It can be seen from Eq. (9.7.36) that F_n is dependent on the renormalization scale μ. Write the dependence of the Feynman amplitude F_n on the scale μ explicitly, then

$$F_n\left(p,g_{\mathrm{r}}(\mu),m_{\mathrm{r}}(\mu),\alpha_{\mathrm{r}}(\mu),\mu\right) = \tilde{G}^{\mathrm{tc}}_{\mathrm{r}n}\left(p,g_{\mathrm{r}}(\mu),m_{\mathrm{r}}(\mu),\alpha_{\mathrm{r}}(\mu),\mu\right) \tag{9.7.37}$$

From Eqs. (9.7.36) and (9.7.37), we can obtain the finite renormalization for F_n between two different renormalization schemes

$$F_n\left(p,g_{\mathrm{r}}(\mu'),m_{\mathrm{r}}(\mu'),\alpha_{\mathrm{r}}(\mu'),\mu'\right) = z_n(\mu',\mu)F_n\left(p,g_{\mathrm{r}}(\mu),m_{\mathrm{r}}(\mu),\alpha_{\mathrm{r}}(\mu),\mu\right) \tag{9.7.38}$$

$$z_n(\mu',\mu) = Z_3^{n/2}(\mu')/Z_3^{n/2}(\mu) \tag{9.7.39}$$

This means that the change of the renormalization scale from $\mu \to \mu'$ corresponds to a finite transformation $z_n(\mu',\mu)$. The set of these transformation $\{z_n(\mu',\mu)\}$ makes up an Abelian group.

The renormalization of Green function can be easily extended to the case of including quark fields. For the case of n_{G} gluon fields and n_{F} quark fields, Eq. (9.7.36) becomes

$$\tilde{G}^{\mathrm{tc}}_{\mathrm{r}n}(p,g_{\mathrm{r}},m_{\mathrm{r}},\alpha_{\mathrm{r}}) = Z_3^{n_{\mathrm{G}}/2}(\mu)Z_2^{n_{\mathrm{F}}/2}(\mu)\tilde{G}^{\mathrm{tc}}_n(p,g,m,\alpha) \tag{9.7.40}$$

or

$$F_n\left(p,g_{\mathrm{r}}(\mu),m_{\mathrm{r}}(\mu),\alpha_{\mathrm{r}}(\mu),\mu\right) = Z_3^{n_{\mathrm{G}}/2}(\mu)Z_2^{n_{\mathrm{F}}/2}(\mu)\tilde{G}^{\mathrm{tc}}_n(p,g,m,\alpha) \tag{9.7.41}$$

where Z_2 is the renormalization constant of quark field, $\psi = Z_2^{1/2}\psi_{\mathrm{r}}$, $n = n_{\mathrm{G}} + n_{\mathrm{F}}$. Meanwhile Eq. (9.7.39) should be extended to

$$z_n(\mu',\mu) = Z_3^{n_{\mathrm{G}}/2}(\mu')Z_2^{n_{\mathrm{F}}/2}(\mu')/[Z_3^{n_{\mathrm{G}}/2}(\mu)Z_2^{n_{\mathrm{F}}/2}(\mu)] \tag{9.7.42}$$

Above we have discussed three Abelian groups resulting from the finite renormalizations: $\{z_g(\mu',\mu)\}$, $\{z_m(\mu',\mu)\}$, $\{z_n(\mu',\mu)\}$, which correspond to scale transformation, and are referred to as the renormalization group [37–39].

9.7.2. *The renormalization group equations*

After renormalization, the parameters in the theory, the coupling constant $g_{\mathrm{r}}(\mu)$, the quark mass $m_{\mathrm{r}}(\mu)$, the gauge fixing parameter $\alpha_{\mathrm{r}}(\mu)$, and the Feynman amplitude $F_n\left(p, g_{\mathrm{r}}(\mu), m_{\mathrm{r}}(\mu), \alpha_{\mathrm{r}}(\mu), \mu\right)$ of particle scattering are all dependent on the renormalization scale. If the renormalization scale is made a small change, then the coupling constant, quark mass, gauge fixing parameter and the Feynman amplitude also slightly change accordingly. This change can generate a series of differential equations. These differential equations that describe the theoretical parameters and Feynman amplitude varying with the renormalization scale are called the renormalization group equation [27, 37–42].

At first we shall derive the renormalization group equations for the coupling constant, quark mass, gauge fixing parameter. Equation (9.7.2) can be written as

$$g_{\mathrm{r}}(\mu) = Z_g^{-1}(\mu) g \tag{9.7.43}$$

$$m_{\mathrm{r}}(\mu) = Z_m^{-1}(\mu) m \tag{9.7.44}$$

$$\alpha_{\mathrm{r}}(\mu) = Z_3^{-1}(\mu) \alpha \tag{9.7.45}$$

We derive the renormalization group equations in the MS (or $\overline{\mathrm{MS}}$) scheme. The coupling constant has a dimension of mass in the $D = 4-2\varepsilon$ dimension. The mass dimension of the coupling constant can be separated as

$$g = g_0 \mu_0^{\varepsilon}, \quad g_{\mathrm{r}} = g_{\mathrm{R}} \mu^{\varepsilon} \tag{9.7.46}$$

where the scale μ_0 of the bare coupling constant is fixed, while the scale μ of g_{r} is variable, which is the renormalization scale in the MS or $\overline{\mathrm{MS}}$ subtraction scheme. With Eq. (9.7.46), Eq. (9.7.43) can be written as

$$g_{\mathrm{R}}(\mu) = \left(\frac{\mu_0}{\mu}\right)^{\varepsilon} Z_g^{-1}(\mu) g_0 \tag{9.7.47}$$

The bare parameters g, m, α are all treated as renormalization scale independent constants

$$\frac{\mathrm{d}g}{\mathrm{d}\mu} = 0, \quad \frac{\mathrm{d}m}{\mathrm{d}\mu} = 0, \quad \frac{\mathrm{d}\alpha}{\mathrm{d}\mu} = 0 \tag{9.7.48}$$

Therefore, we can obtain the differential equations for g_{R}, m_{R} and α_{R} from Eqs. (9.7.47), (9.7.44) and (9.7.45) (here we write m_{r} and α_{r} uniformly as

m_{R} and α_{R}, i.e. $m_{\mathrm{R}} = m_{\mathrm{r}}$, $\alpha_{\mathrm{R}} = \alpha_{\mathrm{r}}$),

$$\mu \frac{\mathrm{d}g_{\mathrm{R}}}{\mathrm{d}\mu} = \beta \tag{9.7.49}$$

$$\mu \frac{\mathrm{d}m_{\mathrm{R}}}{\mathrm{d}\mu} = -m_{\mathrm{R}}\gamma_m \tag{9.7.50}$$

$$\mu \frac{\mathrm{d}\alpha_{\mathrm{R}}}{\mathrm{d}\mu} = -\alpha_{\mathrm{R}}\gamma_\alpha \tag{9.7.51}$$

where the functions β, γ_m and γ_α are

$$\beta = -\varepsilon g_{\mathrm{R}} - \frac{\mu}{Z_g}\frac{\mathrm{d}Z_g}{\mathrm{d}\mu} g_{\mathrm{R}} \tag{9.7.52}$$

$$\gamma_m = \frac{\mu}{Z_m}\frac{\mathrm{d}Z_m}{\mathrm{d}\mu} \tag{9.7.53}$$

$$\gamma_\alpha = \frac{\mu}{Z_3}\frac{\mathrm{d}Z_3}{\mathrm{d}\mu} \tag{9.7.54}$$

Although there are infinite terms in Z_g, Z_m and Z_3, the functions β, γ_m and γ_α defined above are finite functions of μ, because the divergent terms in Z_g, Z_m and Z_3 can cancel between numerator and denominators when taking the limit $\varepsilon \to 0$. The functions of β, γ_m and γ_α partially depend on the scale μ directly, and partially indirectly through $g_{\mathrm{R}}(\mu)$, $m_{\mathrm{R}}(\mu)$ and $\alpha_{\mathrm{R}}(\mu)$. Generally, Z_g, Z_m and Z_3 are simultaneously functions of g_{R} and α_{R}. So, in principle, these three functions are intercorrelated, so are the three equations of (9.7.49)–(9.7.51). The functions β, γ_m and γ_α are all not physical observable, because they are in principle gauge dependent. But functions β and γ_m are gauge independent in MS and $\overline{\mathrm{MS}}$ subtraction scheme, which can be shown by the result of the actual calculation in MS scheme [29, 43, 44]. In MS scheme, the renormalization constant Z_g shows up in the following form:

$$Z_g = 1 + \frac{a_1(g_{\mathrm{R}}, \alpha_{\mathrm{R}})}{\varepsilon} + \frac{a_2(g_{\mathrm{R}}, \alpha_{\mathrm{R}})}{\varepsilon^2} + \cdots \tag{9.7.55}$$

where $a_1(g_{\mathrm{R}}, \alpha_{\mathrm{R}})$, $a_2(g_{\mathrm{R}}, \alpha_{\mathrm{R}})$, $\cdots$ are all function of g_{R} and α_{R} calculated in MS scheme. Since the bare coupling constant is gauge independent, so from Eq. (9.7.47) for fixed μ, μ_0 and ε, we have

$$\frac{\mathrm{d}}{\mathrm{d}\alpha_{\mathrm{R}}}(Z_g g_{\mathrm{R}}) = 0 \tag{9.7.56}$$

Substituting Eq. (9.7.55) into Eq. (9.7.56), and expanding it according to the order of $1/\varepsilon$, we can obtain

$$\frac{\mathrm{d}g_\mathrm{R}}{\mathrm{d}\alpha_\mathrm{R}} + \frac{1}{\varepsilon}\left(a_1 \frac{\mathrm{d}g_\mathrm{R}}{\mathrm{d}\alpha_\mathrm{R}} + \frac{\mathrm{d}a_1}{\mathrm{d}\alpha_\mathrm{R}} g_\mathrm{R}\right) + \frac{1}{\varepsilon^2}\left(a_2 \frac{\mathrm{d}g_\mathrm{R}}{\mathrm{d}\alpha_\mathrm{R}} + \frac{\mathrm{d}a_2}{\mathrm{d}\alpha_\mathrm{R}} g_\mathrm{R}\right) + \cdots = 0 \tag{9.7.57}$$

Because the terms of different order of $1/\varepsilon$ are independent, they must be all zero. Then we have

$$\frac{\mathrm{d}g_\mathrm{R}}{\mathrm{d}\alpha_\mathrm{R}} = 0, \quad \frac{\mathrm{d}a_1}{\mathrm{d}\alpha_\mathrm{R}} = 0, \quad \frac{\mathrm{d}a_2}{\mathrm{d}\alpha_\mathrm{R}} = 0 \tag{9.7.58}$$

This indicates that g_R, a_1, a_2, $\cdots$ are all actually gauge independent. Therefore, we can conclude that β function is gauge independent in MS subtraction scheme. The same conclusion can be also drawn for γ_m function.

It is worth mentioning that the renormalization constant Z_g of the coupling constant is in general gauge dependent in other renormalization schemes. The reason is that, different from Eq. (9.7.55), the result of Z_g has the following form:

$$Z_g = 1 + a_0(g_\mathrm{R}, \alpha_\mathrm{R}) + \frac{a_1(g_\mathrm{R}, \alpha_\mathrm{R})}{\varepsilon} + \frac{a_2(g_\mathrm{R}, \alpha_\mathrm{R})}{\varepsilon^2} + \cdots \tag{9.7.59}$$

in other subtraction schemes. Compared with Eq. (9.7.55), there is an extra term $a_0(g_\mathrm{R}, \alpha_\mathrm{R})$ appearing in Eq. (9.7.59). It is just because of the presence of $a_0(g_\mathrm{R}, \alpha_\mathrm{R})$, the above proof fails.

Next we shall go on to derive the differential equation for the Feynman amplitude F_n from Eq. (9.7.41). We know that unrenormalized n-point Green function $\tilde{G}_n^{\mathrm{tc}}(p, g, m, \alpha)$ is independent on the renormalization scale μ as long as the bare parameters g, m and α are fixed. i.e.

$$\left.\frac{\mathrm{d}}{\mathrm{d}\mu}\tilde{G}_n^{\mathrm{tc}}(p, g, m, \alpha)\right|_{g,m,\alpha} = 0 \tag{9.7.60}$$

where the subscripts g, m, α stands for that g, m and α are fixed. Considering Eq. (9.7.41), we can obtain the following equation:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\mu}\Big[&Z_3^{-n_\mathrm{G}/2}(\mu, g, m, \alpha) Z_2^{-n_\mathrm{F}/2}(\mu, g, m, \alpha) \\ &\cdot F_n\left(p, g_\mathrm{R}(\mu, g, m, \alpha), m_\mathrm{R}(\mu, g, m, \alpha), \alpha_\mathrm{R}(\mu, g, m, \alpha), \mu\right)\Big] = 0\end{aligned} \tag{9.7.61}$$

The dependence of the renormalization constants Z_2, Z_3 and the coupling constant, quark mass and gauge fixing parameter on scale μ, g, m and α

has been written explicitly in Eq. (9.7.61), and $g_{\rm r}$, $m_{\rm r}$, $\alpha_{\rm r}$ been changed to $g_{\rm R}$, $m_{\rm R}$, $\alpha_{\rm R}$. By using the chain rule of differentiation, from Eq. (9.7.61) we can obtain

$$\begin{aligned}&\left.\frac{\partial Z_3^{-n_{\rm G}/2}}{\partial\mu}\right|_{g,m} Z_2^{-n_{\rm F}/2}F_n + Z_3^{-n_{\rm G}/2}\left.\frac{\partial Z_2^{-n_{\rm F}/2}}{\partial\mu}\right|_{g,m} F_n\\
&+Z_3^{-n_{\rm G}/2}Z_2^{-n_{\rm F}/2}\left(\frac{\partial}{\partial\mu}+\frac{\partial g_{\rm R}}{\partial\mu}\frac{\partial}{\partial g_{\rm R}}+\frac{\partial m_{\rm R}}{\partial\mu}\frac{\partial}{\partial m_{\rm R}}\right.\\
&\left.\left.+\frac{\partial\alpha_{\rm R}}{\partial\mu}\frac{\partial}{\partial\alpha_{\rm R}}\right)F_n\right|_{g,m}=0 \qquad (9.7.62)\end{aligned}$$

Equation (9.7.62)can be simplified to be

$$\begin{aligned}&\left[\mu\frac{\partial}{\partial\mu}+\beta(g_{\rm R},\alpha_{\rm R})\frac{\partial}{\partial g_{\rm R}}-\gamma_m(g_{\rm R},\alpha_{\rm R})m_{\rm R}\frac{\partial}{\partial m_{\rm R}}\right.\\
&\left.+\delta(g_{\rm R},\alpha_{\rm R})\frac{\partial}{\partial\alpha_{\rm R}}-n_{\rm G}\gamma_{\rm G}(g_{\rm R},\alpha_{\rm R})-n_{\rm F}\gamma_{\rm F}(g_{\rm R},\alpha_{\rm R})\right]F_n=0 \qquad (9.7.63)\end{aligned}$$

where β, γ_m, δ, $\gamma_{\rm G}$ and $\gamma_{\rm F}$ are called the renormalization group functions. β is also called β function, and $\gamma_{\rm G}$, $\gamma_{\rm F}$ are referred to as anomalous dimensions of the gluon and quark fields, respectively. The definitions of them are as follows:

$$\beta(g_{\rm R},\alpha_{\rm R})=\mu\left.\frac{\partial g_{\rm R}}{\partial\mu}\right|_{g,m,\alpha} \qquad (9.7.64)$$

$$\gamma_m(g_{\rm R},\alpha_{\rm R})=-\left.\frac{\mu}{m_{\rm R}}\frac{\partial m_{\rm R}}{\partial\mu}\right|_{g,m,\alpha} \qquad (9.7.65)$$

$$\delta(g_{\rm R},\alpha_{\rm R})=\mu\left.\frac{\partial\alpha_{\rm R}}{\partial\mu}\right|_{g,m,\alpha}=-2\alpha_{\rm R}\gamma_{\rm G}(g_{\rm R},\alpha_{\rm R}) \qquad (9.7.66)$$

$$\gamma_{\rm G}(g_{\rm R},\alpha_{\rm R})=\left.\frac{\mu}{2Z_3}\frac{\partial Z_3}{\partial\mu}\right|_{g,m,\alpha} \qquad (9.7.67)$$

$$\gamma_{\rm F}(g_{\rm R},\alpha_{\rm R})=\left.\frac{\mu}{2Z_2}\frac{\partial Z_2}{\partial\mu}\right|_{g,m,\alpha} \qquad (9.7.68)$$

We have used $Z_3\alpha_{\rm R}=\alpha$ in the second step of Eq. (9.7.66). Renormalization group functions play an important role in obtaining many basic properties of QCD. It is especially β function, from which the asymptotic freedom of QCD can be derived. This will be discussed later. As mentioned earlier, the renormalized group functions are generally gauge dependent. But β and γ_m functions are all gauge independent in MS and $\overline{\text{MS}}$ schemes.

Equations (9.7.63) and (9.7.64)–(9.7.68) together constitute the renormalization group equations of the Feynman amplitude F_n in MS subtraction scheme. Equation (9.7.63) are called 't Hooft-Weinberg equation [42,45,46].

The renormalization group equation is generally dependent on the subtraction scheme adopted in the renormalization process. In different subtraction schemes, the form of the renormalization group equations will be slightly different, but the general idea of deriving the equation remains unchanged. For example, for the renormalization group equation of scalar field, the renormalized mass of the scalar field is taken to be the pole mass, while the renormalization of the coupling constant and scalar field adopt the off-shell renormalization scheme, the renormalization group equation of Green function obtained in this way is called Gell-Mann–Low equation [38]. Symanzik [40] and Callan [41] adopted on-shell subtraction scheme, the renormalization group equation obtained in such way is called Callan–Symanzik equation. Georgi and Politzer completely adopted off-shell subtraction scheme, the renormalization group equation they obtained is referred to as Georgi-Politzer equation [47].

9.7.3. *The solution of the renormalization group equation*

Now let us solve the renormalization group equation. We discuss this topic in MS subtraction scheme. The equation in MS scheme is 't Hooft-Weinberg equation, as shown in Eq. (9.7.63). Now, for the sake of brevity, we remove the subscript "R" for renormalization in the case without causing confusion. Then 't Hooft–Weinberg equation becomes

$$\left[\mu\frac{\partial}{\partial\mu} + \beta(g,\alpha)\frac{\partial}{\partial g} - \gamma_m(g,\alpha)m\frac{\partial}{\partial m} + \delta(g,\alpha)\frac{\partial}{\partial\alpha} - n_{\text{G}}\gamma_{\text{G}}(g,\alpha)\right.$$
$$\left. - n_{\text{F}}\gamma_{\text{F}}(g,\alpha)\right] F_n(\lambda p, g, m, \alpha, \mu) = 0 \tag{9.7.69}$$

where p stands for the set of all particle momenta, $p = \{p_1, p_2, \ldots, p_n\}$, $n = n_{\text{G}} + n_{\text{F}}$. We also introduce a dimensionless parameter λ representing the momentum scale, which can be conveniently used to express the momentum variation, while the renormalization functions in the equation are given in Eqs. (9.7.64)–(9.7.68).

Suppose that F_n is the truncated Green function of n_{G} gluon fields and n_{F} quark fields, therefore the dimension of this Green function can be calculated to be

$$[F_n] = D - (n_{\mathrm{G}} d_A + n_{\mathrm{F}} d_\psi) \tag{9.7.70}$$

according to Eqs. (9.7.35), (9.7.37) and (9.7.41), where D is the dimension of the space–time, d_A and d_ψ the dimensions of the gluon and quark fields, respectively, and $d_A = (D-2)/2$, $d_\psi = (D-1)/2$. From the dimension of the Feynman amplitude F_n, it can be known that F_n should be in the following general form:

$$F_n(\lambda p, g, m, \alpha, \mu) = \mu^{D-d_0} \phi_n \left(\frac{\lambda p}{\mu}, g, \frac{m}{\mu}, \alpha \right) \tag{9.7.71}$$

where ϕ_n is a dimensionless function, and d_0 is the total dimension of the gluon and quark fields

$$d_0 = n_{\mathrm{G}} d_A + n_{\mathrm{F}} d_\psi \tag{9.7.72}$$

Considering the most general form of F_n, i.e. Eq. (9.7.71), we can know that the following differential equation for F_n holds:

$$\left(\mu \frac{\partial}{\partial \mu} + m \frac{\partial}{\partial m} + \lambda \frac{\partial}{\partial \lambda} - D + d_0 \right) F_n(\lambda p, g, m, \alpha, \mu) = 0 \tag{9.7.73}$$

Subtracting Eq. (9.7.73) from Eq. (9.7.69), we obtain

$$\begin{aligned} \Bigg[-\lambda \frac{\partial}{\partial \lambda} + \beta(g,\alpha) \frac{\partial}{\partial g} - (1 + \gamma_m(g,\alpha))\, m \frac{\partial}{\partial m} + \delta(g,\alpha) \frac{\partial}{\partial \alpha} + D - d_0 \\ - n_{\mathrm{G}} \gamma_{\mathrm{G}} - n_{\mathrm{F}} \gamma_{\mathrm{F}} \Bigg] F_n(\lambda p, g, m, \alpha, \mu) = 0 \end{aligned} \tag{9.7.74}$$

Introducing

$$t = -\ln \lambda, \quad \omega_n(g,\alpha) = D - d_0 - n_{\mathrm{G}} \gamma_{\mathrm{G}} - n_{\mathrm{F}} \gamma_{\mathrm{F}} \tag{9.7.75}$$

Equation (9.7.74) can be rewritten as

$$\begin{aligned} \Bigg[\frac{\partial}{\partial t} + \beta(g,\alpha) \frac{\partial}{\partial g} - (1 + \gamma_m(g,\alpha))\, m \frac{\partial}{\partial m} \\ + \delta(g,\alpha) \frac{\partial}{\partial \alpha} + \omega_n(g,\alpha) \Bigg] F_n(\lambda p, g, m, \alpha, \mu) = 0 \end{aligned} \tag{9.7.76}$$

Equation (9.7.76) is a homogeneous linear partial differential equation of first order. The renormalization functions β, γ_m and δ all do not depend on the scale μ directly. Their dependences on μ are through $g(\mu)$ and $\alpha(\mu)$ indirectly. This makes it easy to solve this equation. We can treat this equation as an equation about the variable t, i.e. we try to find the solution of the Feynman amplitude varying with the momentum. It is worth noting that m, g and α in Eq. (9.7.76) are all functions of the scale parameter μ, and independent of t. We introduce the running coupling constant $\bar{g}(t)$, the running mass $\bar{m}(t)$ and the running gauge fixing parameter $\bar{\alpha}(t)$, which are all functions of t. They are the renormalized functions defined at the scale μ/λ, and they are defined to obey the following equations:

$$\begin{aligned} \frac{\mu}{\lambda}\frac{\mathrm{d}\bar{g}}{\mathrm{d}\left(\mu/\lambda\right)} &= \beta(\bar{g},\bar{\alpha}) \\ \frac{\mu}{\lambda}\frac{1}{\bar{m}}\frac{\mathrm{d}\bar{m}}{\mathrm{d}\left(\mu/\lambda\right)} &= -1-\gamma_m(\bar{g},\bar{\alpha}) \\ \frac{\mu}{\lambda}\frac{\mathrm{d}\bar{\alpha}}{\mathrm{d}\left(\mu/\lambda\right)} &= \delta(\bar{g},\bar{\alpha}) \end{aligned} \tag{9.7.77}$$

Viewing λ as a variable can result in

$$\frac{\mathrm{d}\bar{g}}{\mathrm{d}t} = \beta(\bar{g},\bar{\alpha}), \quad \frac{1}{\bar{m}}\frac{\mathrm{d}\bar{m}}{\mathrm{d}t} = -1-\gamma_m(\bar{g},\bar{\alpha}), \quad \frac{\mathrm{d}\bar{\alpha}}{\mathrm{d}t} = \delta(\bar{g},\bar{\alpha}) \tag{9.7.78}$$

Replacing all the g, m and α in Eq. (9.7.76) by the relevant running quantities $\bar{g}$, $\bar{m}$ and $\bar{\alpha}$, then according to Eq. (9.7.78), one can transform Eq. (9.7.76) to be a total differential equation

$$\left[\frac{\mathrm{d}}{\mathrm{d}t} + \omega_n(\bar{g},\bar{\alpha})\right] F_n(\mathrm{e}^{-t}p,\bar{g},\bar{m},\bar{\alpha},\mu) = 0 \tag{9.7.79}$$

It can be seen from Eq. (9.7.77) that all the $\bar{g}$, $\bar{m}$ and $\bar{\alpha}$ go back to g, m and α, respectively, as $\lambda = 1$ or $t = 0$, i.e.

$$\bar{g}(0) = g, \quad \bar{m}(0) = m, \quad \bar{\alpha}(0) = \alpha \tag{9.7.80}$$

Taking the above equation as the initial condition of running parameters, we can obtain

$$F_n\left(\mathrm{e}^{-t}p,\bar{g}(t),\bar{m}(t),\bar{\alpha}(t),\mu\right) = F_n(p,g,m,\alpha,\mu)\mathrm{e}^{-\int_0^t \mathrm{d}t'\omega_n\left(\bar{g}(t'),\bar{\alpha}(t')\right)} \tag{9.7.81}$$

by integrating Eq. (9.7.79). This is the general solution to 't Hooft–Weinberg equation, by which we can discuss the behavior of Feynman amplitude in the large momentum limit. Doing the substitution $\mathrm{e}^{-t}p \to p$, and changing the form of Eq. (9.7.81) slightly, we can obtain

$$\begin{aligned} F_n\left(\mathrm{e}^t p, g, m, \alpha, \mu\right) = {} & F_n(p, \bar{g}(t), \bar{m}(t), \bar{\alpha}(t), \mu) \\ & \times \exp\left[(D-d_0)t - \int_0^t \mathrm{d}t' \left(n_{\mathrm{G}}\gamma_{\mathrm{G}}(\bar{g}(t'), \bar{\alpha}(t'))\right.\right. \\ & \left.\left. + n_{\mathrm{F}}\gamma_{\mathrm{F}}\left(\bar{g}(t'), \bar{\alpha}(t')\right)\right)\right] \end{aligned} \tag{9.7.82}$$

which gives the Feynman amplitude varying with the particle momentum, and t changes from $-\infty$ to $+\infty$. $t \to +\infty$ is referred to as ultraviolet limit, $t \to -\infty$ the infrared limit. The running coupling constant, running mass and running gauge parameter are determined by the solutions of Eq. (9.7.78). Their integral forms are as follows:

$$t = \int_g^{\bar{g}(t)} \frac{\mathrm{d}g'}{\beta(g')} \tag{9.7.83}$$

$$\bar{m}(t) = m\mathrm{e}^{-t-\int_0^t \mathrm{d}t' \gamma_m\left(\bar{g}(t')\right)} \tag{9.7.84}$$

$$t = \int_\alpha^{\bar{\alpha}(t)} \frac{\mathrm{d}\alpha'}{\delta\left(\bar{g}(t), \alpha'\right)} \tag{9.7.85}$$

It has been considered in Eqs. (9.7.83) and (9.7.84) that both β and γ_m functions are gauge independent in MS scheme. Their calculated result in QCD will be discussed later.

9.8. Asymptotic Freedom

It can be seen from the solution of the renormalization group equation that the renormalization group functions determine the asymptotic behavior of the Feynman amplitude varying with momentum. The renormalization group functions also determine the asymptotic behavior of the running coupling constant, the running mass and the gauge fixing parameter varying with momentum. At first, let us discuss the running coupling constant $\bar{g}(t)$, especially for the behavior of $\bar{g}(t)$ at $t \to \infty$. Since β is calculated in the perturbation theory of QCD, and it has the form of $\beta(g) = bg^n + O(g^{n+1})$, where $n \geq 1$, so we have $\beta(0) = 0$. Let us discuss it in the following for several cases.

(1) If there is no zero for $\beta(\bar{g}))$ except at $\bar{g} = 0$, in addition, $\beta(\bar{g})$ is all positive in the whole range of g, as shown in Fig. 9.15(a), then it can be

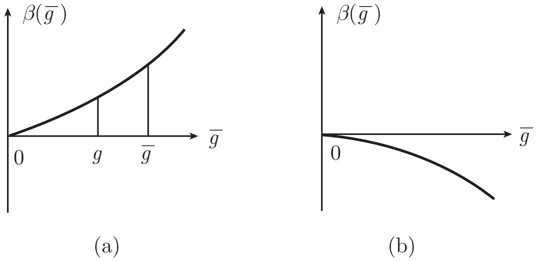


Fig. 9.15. β function varying with $\bar{g}$: (a) the case $\beta(\bar{g})$ is a monotonically increasing function of $\bar{g}$, (b) the case $\beta(\bar{g})$ is a monotonically decreasing function of $\bar{g}$.

known from Eq. (9.7.83) that $\bar{g}(t)$ should have the following asymptotic behavior:

$$\begin{cases} \bar{g}(t) \to \infty, & t \to +\infty \\ \bar{g}(t) \to 0, & t \to -\infty \end{cases} \tag{9.8.1}$$

If $\beta(\bar{g})$ has no zero except $\beta(0) = 0$, and is negative as shown in Fig. 9.15(b), then $\bar{g}(t)$ behaves as

$$\begin{cases} \bar{g}(t) \to 0, & t \to +\infty \\ \bar{g}(t) \to \infty, & t \to -\infty \end{cases} \tag{9.8.2}$$

(2) If β function has a zero at $\bar{g} = g_c$, i.e. $\beta(g_c) = 0$ except $\beta(0) = 0$, and the derivative of β at $\bar{g} = g_c$ is negative $\beta'(g_c) < 0$ as shown in Fig. 9.16(a), then

$$\bar{g}(t) \to g_c, \quad \text{as} \quad t \to +\infty \tag{9.8.3}$$

If $\beta'(g_c) > 0$, as shown in Fig. 9.16(b), then

$$\bar{g}(t) \to g_c, \quad \text{as} \quad t \to -\infty \tag{9.8.4}$$

Case (1) can be regarded as a special case of (2), so let us just discuss case (2). Since the scale of momentum increases as t increases, the limit $t \to \infty$ is called the ultraviolet (UV) limit. Meanwhile, the limit $t \to -\infty$ is called the infrared (IR) limit. The limit value of g_c of $\bar{g}(t)$ corresponding to $t \to +\infty$ is called the ultraviolet fixed point, while the limit value of $\bar{g}(t)$ corresponding to $t \to -\infty$ is known as the infrared fixed point.

As the coupling constant is near the limiting point $g_c = 0$, its value is close to zero. In this case, the interaction is weak, and the theory tends to

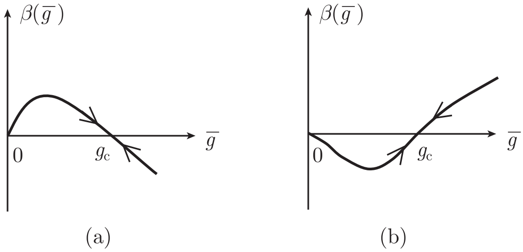


Fig. 9.16. β function varying with $\bar{g}$. There is another zero at $\bar{g} = g_c$ for β function, $\beta(g_c) = 0$ except the point $\bar{g} = 0$. (a) The case for $\beta'(g_c) < 0$, where the arrow points to the moving direction of $\bar{g}(t)$ as $t \to \infty$; (b) the case for $\beta'(g_c) > 0$, where the arrow points to the direction of $\bar{g}(t)$ as $t \to -\infty$.

be free field theory. If $g_c = 0$ is the ultraviolet fixed point, then the theory is asymptotically free.

Let us take an infrared and an ultraviolet fixed points as two examples. The β function of QED calculated up to one-loop order is in the following form:

$$\beta(\bar{g}) = \beta_0 \bar{g}^3 + O(\bar{g}^5), \quad \beta_0 > 0 \tag{9.8.5}$$

So, $\bar{g} = 0$ is the infrared fixed point of QED. Substituting Eq. (9.8.5) into Eq. (9.7.83) and integrating, we can obtain

$$\bar{g}(t)^2 = \frac{g^2}{1 - 2\beta_0 g^2 t} \tag{9.8.6}$$

As $t \to -\infty$, $\bar{g}(t) \to 0$, so it is indeed the infrared fixed point. Let us consider the region $t > 0$. When $t = \frac{1}{2\beta_0 g^2}$, i.e. when the momentum scale is at $e^t = \exp\left(\frac{1}{2\beta_0 g^2}\right)$, the coupling constant approaches to infinity, so perturbation theory fails. This means the existence of a pole on the scale of momentum, which is known as the Landau pole [48, 49].

For the following β function

$$\beta(\bar{g}) = -\beta_0 \bar{g}^3 + O(\bar{g}^5), \quad \beta_0 > 0 \tag{9.8.7}$$

substituting it into Eq. (9.7.83) and doing the integration, we can obtain

$$\bar{g}(t)^2 = \frac{g^2}{1 + 2\beta_0 g^2 t} \tag{9.8.8}$$

As $t \to +\infty$, there is $\bar{g}(t) \to 0$, so this is the ultraviolet fixed point.

Now let us calculate the β function in QCD. We go on omitting the subscript "R" for the renormalized quantities. At first, we discuss the calculation of one-loop order. It is known from Eq. (9.7.52), the definition of the β function (note that the subscript "R" will be omitted here), that one needs to know the renormalization constant Z_g of the coupling constant to calculate the β function. We know from Eq. (9.5.33) that there are totally four methods to calculate Z_g. But the calculated results through these four ways are all the same according to Slavnov–Taylor identity in Eq. (9.5.35). Let us calculate Z_g by

$$Z_g = \tilde{Z}_1/(\tilde{Z}_3 Z_3^{1/2}) \tag{9.8.9}$$

We need to calculate the ghost–gluon interaction vertex, the ghost self-energy, the gluon self-energy to one-loop order, so that we can obtain $\tilde{Z}_1$, $\tilde{Z}_3$, $Z_3^{1/2}$, which have been given in Eqs. (9.5.74), (9.5.66) and (9.5.64). Removing the subscript "R" or "r", the results are as follows:

$$\tilde{Z}_1 = 1 - \frac{g^2}{(4\pi)^2} C_{\mathrm{G}} \frac{\alpha}{2}\frac{1}{\varepsilon} + O(g^4) \tag{9.8.10}$$

$$\tilde{Z}_3 = 1 + \frac{g^2}{(4\pi)^2} C_{\mathrm{G}} \frac{3-\alpha}{4}\frac{1}{\varepsilon} + O(g^4) \tag{9.8.11}$$

$$Z_3 = 1 + \frac{g^2}{(4\pi)^2}\left[\frac{1}{2}C_{\mathrm{G}}\left(\frac{13}{3}-\alpha\right) - \frac{4}{3}T_{\mathrm{R}}N_{\mathrm{f}}\right]\frac{1}{\varepsilon} + O(g^4) \tag{9.8.12}$$

Substituting the above three results into Eq. (9.8.9) results in

$$Z_g = 1 - \frac{g^2}{(4\pi)^2}\frac{1}{\varepsilon}\frac{1}{6}(11C_{\mathrm{G}} - 4T_{\mathrm{R}}N_f) + O(g^4) \tag{9.8.13}$$

Just as mentioned in Sec. 9.7.2, the result of Z_g does not depend on the gauge parameter α because the dependence canceled in the calculation of Z_g. Now we substitute Eq. (9.8.13) into Eq. (9.7.52), then we obtain

$$\begin{aligned}\beta(g) &= -\varepsilon g - \frac{\mu}{Z_g}\frac{\mathrm{d}Z_g}{\mathrm{d}\mu} g \\ &= -\varepsilon g + \frac{11C_{\mathrm{G}} - 4T_{\mathrm{R}}N_{\mathrm{f}}}{3}\frac{g^2}{(4\pi)^2}\frac{1}{\varepsilon}\beta(g) + O(g^4)\end{aligned} \tag{9.8.14}$$

Iterating Eq. (9.8.14) itself into $\beta(g)$ in the right-hand side, remaining up to the g^3 term and taking $\varepsilon \to 0$, we obtain

$$\beta(g) = -\frac{1}{(4\pi)^2}\frac{11C_{\mathrm{G}} - 4T_{\mathrm{R}}N_{\mathrm{f}}}{3} g^3 + O(g^5) \tag{9.8.15}$$

This result is in the form

$$\beta(\bar{g}) = -\beta_0 g^3 + O(g^5) \tag{9.8.16}$$

$$\beta_0 = \frac{1}{(4\pi)^2}\frac{11C_{\mathrm{G}} - 4T_{\mathrm{R}}N_{\mathrm{f}}}{3} \tag{9.8.17}$$

As long as $\beta_0 > 0$, QCD should have the property of asymptotic freedom, for which the condition is $11C_{\mathrm{G}} - 4T_{\mathrm{R}}N_{\mathrm{f}} > 0$. For SU(3) group, $C_{\mathrm{G}} = 3$, $T_{\mathrm{R}} = 1/2$, so the condition of asymptotic freedom is

$$N_{\mathrm{f}} < 33/2 \tag{9.8.18}$$

This means that QCD has the property of asymptotic freedom as long as the number of quark flavor is less than 16. The result of asymptotic freedom can be attributed to the three-gluon coupling in the Lagrangian of QCD, which is the self-interaction term of the gluon fields. Since the self-interaction of the gauge field is the characteristic of Yang–Mills theory, asymptotic freedom is the essential characteristic of Yang–Mills theory.

The other renormalization group functions γ_m, δ, γ_{G} and γ_{F} can be directly calculated according to their definitions and the renormalization constants given in Sec. 9.5.4, which will not be discussed in detail here.

The calculation of renormalization group functions beyond one-loop order becomes very difficult. Now the calculation for the β function has been carried out to the four-loop order in the MS subtraction scheme. The first complete two-loop calculation has been done by Jones [50] and Caswell [51]. The three and four-loop calculations are very difficult, which were completed in 1980 and 1997 [52, 53]. The result of β function up to three-loop is

$$\beta(g) = -\beta_0 g^3 - \beta_1 g^5 - \beta_2 g^7 + O(g^9) \tag{9.8.19}$$

where β_0 is given in Eq. (9.8.17), and

$$\beta_1 = \frac{1}{(4\pi)^4}\left[\frac{34}{3}C_{\mathrm{G}}^2 - 4\left(\frac{5}{3}C_{\mathrm{G}} + C_{\mathrm{F}}\right)T_{\mathrm{R}}N_{\mathrm{f}}\right] \tag{9.8.20}$$

$$\begin{aligned}\beta_2 = \frac{1}{(4\pi)^6}&\left[\frac{2857}{54}\,C_{\mathrm{G}}^3 - \left(\frac{1415}{27}C_{\mathrm{G}}^2 + \frac{205}{9}C_{\mathrm{G}}C_{\mathrm{F}} - 2C_{\mathrm{F}}^2\right)T_{\mathrm{R}}N_{\mathrm{f}}\right.\\ &\left.+\left(\frac{158}{27}C_{\mathrm{G}} + \frac{44}{9}C_{\mathrm{F}}\right)T_{\mathrm{R}}^2N_{\mathrm{f}}^2\right]\end{aligned} \tag{9.8.21}$$

For the color SU(3) group, the results of β_0, β_1 and β_2 are

$$\beta_0 = \frac{1}{(4\pi)^2}\left(11 - \frac{2}{3}N_{\mathrm{f}}\right) \tag{9.8.22}$$

$$\beta_1 = \frac{1}{(4\pi)^4}\left(102 - \frac{38}{3}N_{\mathrm{f}}\right) \tag{9.8.23}$$

$$\beta_2 = \frac{1}{(4\pi)^6}\left(\frac{2857}{2} - \frac{5033}{18}N_{\mathrm{f}} + \frac{325}{54}N_{\mathrm{f}}^2\right) \tag{9.8.24}$$

Substituting Eq. (9.8.19) into Eq. (9.7.83), one can obtain the integral equation of the running coupling constant up to three-loop

$$t = -\int_g^{\bar{g}(t)} \frac{\mathrm{d}g'}{g'^3\left(\beta_0 + \beta_1 g'^2 + \beta_2 g'^4 + O(g'^6)\right)} \tag{9.8.25}$$

For convenience later, let us take the scale of momentum as

$$\mathrm{e}^t = \sqrt{q^2}/\mu \tag{9.8.26}$$

Here q is the typical momentum transfer in the scattering process, which is time-like, i.e. $q^2 > 0$, while μ is a fixed scale. As $t = 0$, i.e. $\sqrt{q^2} = \mu$, $\bar{g}(0) = g$, that means g is the coupling constant relevant to the scale μ. Equation (9.8.25) can be expanded with respect to g' and the expansion series can be truncated to a finite terms, if one can select an appropriate scale μ, at which g is small enough and $\bar{g}(t) < g$, so that g' is also small enough. Up to one-loop order, we can obtain from Eq. (9.8.25)

$$t = \frac{1}{2\beta_0}\left(\frac{1}{\bar{g}^2} - \frac{1}{g^2}\right) \tag{9.8.27}$$

Then the running coupling constant can be solved to be

$$\begin{aligned}\bar{g}^2 &= \frac{g^2}{1 + 2\beta_0 g^2 t} = \frac{g^2}{1 + 2\beta_0 g^2 \ln \frac{\sqrt{q^2}}{\mu}} \\ &= \frac{1}{\beta_0\left(\frac{1}{\beta_0 g^2} + \ln \frac{q^2}{\mu^2}\right)} = \frac{1}{\beta_0\left(\ln \mathrm{e}^{1/(\beta_0 g^2)} + \ln \frac{q^2}{\mu^2}\right)} \\ &= \frac{1}{\beta_0 \ln \frac{q^2}{\mu^2 \mathrm{e}^{-1/(\beta_0 g^2)}}}\end{aligned} \tag{9.8.28}$$

Define the parameter as

$$\Lambda = \mu \mathrm{e}^{-\frac{1}{2\beta_0 g^2}} \tag{9.8.29}$$

Then Eq. (9.8.28) can be simply expressed as

$$\bar{g}^2 = \frac{1}{\beta_0 \ln \frac{q^2}{\Lambda^2}} \tag{9.8.30}$$

It is seen from the definition of Λ in Eq. (9.8.29) that, Λ is superficially a parameter depending on the scale μ, actually it is a scale-independent constant. The dependence of g^2 on μ can cancel the μ-dependence of Λ in Eq. (9.8.29), since g^2 is the coupling constant relevant to scale μ. The proof is given below.

Suppose that there is another new scale μ', at which the relevant coupling constant is $g(\mu')$. Then we can define a new parameter according to Eq. (9.8.29)

$$\Lambda' = \mu' \mathrm{e}^{-\frac{1}{2\beta_0 \bar{g}^2(\mu')}} \tag{9.8.31}$$

According to Eq. (9.8.30), replacing q^2 to scale μ', one obtain

$$\bar{g}^2(\mu') = \frac{1}{\beta_0 \ln \frac{\mu'^2}{\Lambda^2}} \tag{9.8.32}$$

Substituting Eq. (9.8.32) into Eq. (9.8.31), one can obtain after a simple treatment

$$\Lambda' = \Lambda \tag{9.8.33}$$

This means that the parameter Λ defined in Eq. (9.8.29) is a scale μ independent parameter.

The parameter Λ is usually called the QCD scale parameter, which is the only parameter in QCD except for the quark mass. Its value needs to be determined by comparing the theoretical prediction of QCD to the scattering cross section and the experimental measurement.

If including the two-loop contribution in QCD, Eq. (9.8.30) can be further corrected. Approximating Eq. (9.8.25) to the term of β_1, and integrating it, we can obtain

$$t = \frac{1}{2\beta_0}\left[\frac{1}{\bar{g}^2} - \frac{1}{g^2} + \frac{\beta_1}{\beta_0} \ln \frac{\bar{g}^2(\beta_0 + \beta_1 g^2)}{g^2(\beta_0 + \beta_1 \bar{g}^2)}\right] \tag{9.8.34}$$

Separating the terms of $\bar{g}$ in Eq. (9.8.34) on one side, and using $t = \ln(\sqrt{q^2}/\mu)$, we can obtain

$$\frac{1}{\bar{g}^2} + \frac{\beta_1}{\beta_0} \ln \frac{\beta_0 \bar{g}^2}{1 + \frac{\beta_1}{\beta_0}\bar{g}^2} = \beta_0 \ln \frac{q^2}{\mu^2} + \frac{1}{g^2} - \frac{\beta_1}{\beta_0} \ln \frac{1 + \beta_1 g^2/\beta_0}{\beta_0 g^2} \tag{9.8.35}$$

The right-hand side of Eq. (9.8.35) can be further treated, and finally one can obtain

$$\frac{1}{\bar{g}^2}+\frac{\beta_1}{\beta_0}\ln\frac{\beta_0\bar{g}^2}{1+\frac{\beta_1}{\beta_0}\bar{g}^2}=\beta_0\ln\frac{q^2}{\Lambda^2} \tag{9.8.36}$$

where Λ is defined as

$$\Lambda=\mu\mathrm{e}^{-1/(2\beta_0 g^2)}\left(\frac{1+\beta_1 g^2/\beta_0}{\beta_0 g^2}\right)^{\beta_1/(2\beta_0^2)} \tag{9.8.37}$$

In the case $\beta_1 = 0$, Eq. (9.8.37) comes back to the one-loop result of Λ given in Eq. (9.8.29). Equation (9.8.36) can be solved by iteration. For the case of $q^2 \gg \Lambda^2$, the solution by iteration is

$$\bar{g}^2=\frac{1}{\beta_0\ln\frac{q^2}{\Lambda^2}}\left(1-\frac{\beta_1}{\beta_0^2}\frac{\ln\ln\frac{q^2}{\Lambda^2}}{\ln\frac{q^2}{\Lambda^2}}+\cdots\right) \tag{9.8.38}$$

The running coupling constant given in Eq. (9.8.38) decreases as the momentum transfer q^2 increasing, where the first term is the one-loop contribution, and the second term the two-loop contribution. The perturbative expansion can apply as q^2 being large enough, then $\bar{g}^2$ being effectively small, and the theory goes to free field theory as $q^2 \to \infty$ and $\bar{g}^2 \to 0$.

Asymptotic freedom of QCD makes it possible to calculate the scattering of quarks and gluons with perturbation method under the condition of large momentum transfer. But there is no free quarks and gluons in the nature and quarks and gluons are always bound within hadrons. The dynamics of binding quarks and gluons together is not dominated by high momentum transfer, so the perturbation calculation with quarks and gluons as free particles cannot be directly applied to the scattering processes where hadrons are involved as initial and final states. Even so, asymptotic freedom ensures that quarks and gluons within small distance tend to be quasi-free particles, and the calculation of treating quarks and gluons as free particles are not completely meaningless. The factorization method based on the operator product expansion (OPE) can separate the interaction of small and long distances apart, which enables the perturbation calculation for the small distance scattering to be reasonably applied to the scattering process of hadrons.

9.9. Structure Function, Parton Distribution Function and the Application in Perturbative QCD

We have introduced the structure function of nucleon for e–p deep inelastic scattering and simply given the relation between the structure function and the parton distribution function of nucleon in Sec. 6.5–6.7. Here in this section we will discuss further the structure function for lepton–nucleon deep inelastic scattering and the relation between the structure function and the parton distribution function. Finally, examples are given to illustrate their application in the calculation of perturbation QCD.

9.9.1. *Structural function of deep inelastic scattering process*

Lepton–nucleon deep inelastic scattering process $lN \to l'\mathrm{X}$ can be shown in Fig. 9.17, where k and k' are the four-momentum of the in and out-going leptons, respectively. P is the momentum of the nucleon and M the nucleon mass. W is the invariant mass of the final hadronic state X. The exchanged particle can be γ, $W^{\pm}$ and Z, the momentum of which is $q = k - k'$. Let's start by listing some useful invariants:

$$\nu = q \cdot p/M = E - E'$$

which is the energy lost by the lepton in the rest frame of the nucleon, and E and E' are energies of the lepton in the initial and final states, respectively;

$$Q^2 = -q^2 = 2(EE' - \vec{k} \cdot \vec{k'}) - m_l^2 - m_{l'}^2$$
$$x = Q^2/(2M\nu)$$

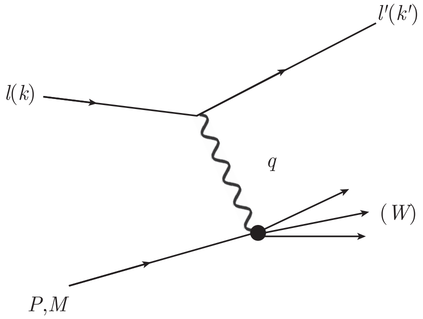


Fig. 9.17. Lepton–nucleon inelastic scattering.

In the parton model, x is the fraction of nucleon momentum carried by quarks involved in collisions;

$$y = q \cdot P/k \cdot P = \nu/E$$

is the fraction of lepton energy loss in the rest frame of the nucleon;

$$W^2 = (P+q)^2 = M^2 + 2M\nu - Q^2$$

is the invariant mass squared of the X system in the final state;

$$s = (k+P)^2 = \frac{Q^2}{xy} + M^2 + m_l^2$$

is the energy squared of the lepton–nucleon system in the center-of-mass frame. The condition of the deep inelastic scattering is $Q^2 \gg M^2$.

In the lowest order of perturbation theory, the differential cross section of the deep inelastic scattering process for a polarized lepton and polarized nucleon shown in Fig. 9.17 can be expressed as

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}x\mathrm{d}y} = \frac{2\pi y\alpha^2}{Q^4}\sum_j \eta_j L_j^{\mu\nu} W_{\mu\nu}^j \tag{9.9.1}$$

For the neutral current: the sum over j is taken as $j = \gamma$ (exchanging photon), $j = Z$ (exchanging Z boson), $j = \gamma Z$ (the interference of exchanging γ and Z);

For the charged current: the sum over j is only for $j = W$ (exchanging W boson);

For the ingoing lepton with charge $e = \pm 1$ and helicity $\lambda = \pm 1$, the leptonic tensor in Eq. (9.9.1) is

$$\begin{aligned} L_{\mu\nu}^{\gamma} &= 2(k_\mu k'_\nu + k'_\mu k_\nu - k \cdot k' g_{\mu\nu} - \mathrm{i}\lambda\varepsilon_{\mu\nu\alpha\beta}k^\alpha k'^\beta) \\ L_{\mu\nu}^{\gamma Z} &= (g_{\mathrm{V}}^{\mathrm{e}} + e\lambda g_{\mathrm{A}}^{\mathrm{e}})L_{\mu\nu}^{\gamma}, \quad L_{\mu\nu}^{Z} = (g_{\mathrm{V}}^{\mathrm{e}} + e\lambda g_{\mathrm{A}}^{\mathrm{e}})^2 L_{\mu\nu}^{\gamma} \\ L_{\mu\nu}^{W} &= (1+e\lambda)^2 L_{\mu\nu}^{\gamma}, \quad g_{\mathrm{V}}^{\mathrm{e}} = -\frac{1}{2} + 2\sin^2\theta_W, \quad g_{\mathrm{A}}^{\mathrm{e}} = -\frac{1}{2} \end{aligned} \tag{9.9.2}$$

where

$$\begin{aligned} &\eta_\gamma = 1, \quad \eta_{\gamma Z} = \left(\frac{G_{\mathrm{F}}M_Z^2}{2\sqrt{2}\pi\alpha}\right)\frac{Q^2}{Q^2+M_Z^2} \\ &\eta_Z = \eta_{\gamma Z}^2, \quad \eta_W = \frac{1}{2}\left(\frac{G_{\mathrm{F}}M_W^2}{4\pi\alpha}\frac{Q^2}{Q^2+M_W^2}\right)^2 \end{aligned} \tag{9.9.3}$$

The hadronic tensor is

$$W_{\mu\nu} = \frac{1}{4\pi} \int \mathrm{d}^4 z \mathrm{e}^{\mathrm{i}q\cdot z} \langle P, S | [J_\mu^+(z), J_\nu(0)] \, | P, S \rangle \tag{9.9.4}$$

where S is the spin four-vector of the nucleon, and $S^2 = -M^2$, $S \cdot P = 0$. It has been obtained from the analysis that, for the scattering of the polarized lepton and nucleon [54], we have

$$\begin{aligned} W_{\mu\nu} = {} & \left(-g_{\mu\nu} + \frac{q_\mu q_\nu}{q^2} \right) F_1(x, Q^2) + \frac{\hat{P}_\mu \hat{P}_\nu}{P \cdot q} F_2(x, Q^2) \\ & - \mathrm{i}\varepsilon_{\mu\nu\alpha\beta} \frac{q^\alpha P^\beta}{2P \cdot q} F_3(x, Q^2) \\ & + \mathrm{i}\varepsilon_{\mu\nu\alpha\beta} \frac{q^\alpha}{P \cdot q} \left[S^\beta g_1(x, Q^2) + \left(S^\beta - \frac{S \cdot q}{P \cdot q} P^\beta \right) g_2(x, Q^2) \right] \\ & + \frac{1}{P \cdot q} \left[\frac{1}{2} (\hat{P}_\mu \hat{S}_\nu + \hat{S}_\mu \hat{P}_\nu) - \frac{S \cdot q}{P \cdot q} \hat{P}_\mu \hat{P}_\nu \right] g_3(x, Q^2) \\ & + \frac{S \cdot q}{P \cdot q} \left[\frac{\hat{P}_\mu \hat{P}_\nu}{P \cdot q} g_4(x, Q^2) + \left(-g_{\mu\nu} + \frac{q_\mu q_\nu}{q^2} \right) g_5(x, Q^2) \right] \end{aligned} \tag{9.9.5}$$

where

$$\hat{P}_\mu = P_\mu - \frac{P \cdot q}{q^2} q_\mu, \quad \hat{S}_\mu = S_\mu - \frac{S \cdot q}{q^2} q_\mu \tag{9.9.6}$$

If the nucleon is unpolarized, the cross section of the deep inelastic scattering process of the polarized lepton on the unpolarized nucleon induced by the neutral or charged current is

$$\begin{aligned} \frac{\mathrm{d}^2\sigma}{\mathrm{d}x\mathrm{d}y} = {} & \frac{4\pi\alpha^2}{xyQ^2} \eta^i \left\{ \left(1 - y - \frac{x^2 y^2 M^2}{Q^2} \right) F_2^i(x, Q^2) + y^2 x F_1^i(x, Q^2) \right. \\ & \left. \mp \left(y - \frac{y^2}{2} \right) x F_3^i(x, Q^2) \right\} \end{aligned} \tag{9.9.7}$$

where i = NC (neutral current of eN → eX), i = CC (charged current eN → νX or νN → eX). For a process of neutrino scattering, the relevant leptonic tension $L_{\mu\nu}^W$ in Eq. (9.9.2) is still correct, with only e and λ corresponding to the charged outcoming lepton.

The negative sign in the last term of Eq. (9.9.7) corresponds to the incoming of e^+ or $\bar{\nu}$, and positive sign corresponds to the incoming of e^-

or ν. The coefficient η^i is

$$\begin{aligned}\eta^{\mathrm{NC}} &= 1 \quad (\text{for unpolarized}\, \mathrm{e}^{\pm}) \\ \eta^{\mathrm{CC}} &= (1 \pm \lambda)^2 \eta_W\end{aligned} \tag{9.9.8}$$

The sign ± in the second term of Eq. (9.9.8) is for the incoming of $l^{\pm}$, λ is the helicity of the incoming lepton, and η_W can be seen in Eq. (9.9.3).

For the case of neutrino incoming, $\eta^{\mathrm{CC}} = 4\eta_W$.

For the structure function generated by the charged current initiated by W exchange, there are

$$F_1^{\mathrm{CC}} = F_1^W, \quad F_2^{\mathrm{CC}} = F_2^W, \quad xF_3^{\mathrm{CC}} = xF_3^W \tag{9.9.9}$$

For the relevant NC structure functions in $\mathrm{e}^{\pm}\mathrm{N} \to \mathrm{e}^{\pm}\mathrm{X}$, which are combinations of F_2^{γ}, $F_2^{\gamma Z}$, F_2^{Z}

$$\begin{aligned}F_2^{\mathrm{NC}} &= F_2^{\gamma} - (g_{\mathrm{V}}^{\mathrm{e}} \pm \lambda g_{\mathrm{A}}^{\mathrm{e}})\eta_{\gamma Z} F_2^{\gamma Z} \\ &\quad + [(g_{\mathrm{V}}^{\mathrm{e}})^2 + (g_{\mathrm{A}}^{\mathrm{e}})^2 \pm 2\lambda g_{\mathrm{V}}^{\mathrm{e}} g_{\mathrm{A}}^{\mathrm{e}}] \cdot \eta_Z F_2^Z \end{aligned} \tag{9.9.10}$$

$$\begin{aligned}xF_3^{\mathrm{NC}} &= -(g_{\mathrm{A}}^{\mathrm{e}} \pm \lambda g_{\mathrm{V}}^{\mathrm{e}})\eta_{\gamma Z} xF_3^{\gamma Z} + [2 g_{\mathrm{V}}^{\mathrm{e}} g_{\mathrm{A}}^{\mathrm{e}} \pm \lambda((g_{\mathrm{V}}^{\mathrm{e}})^2 \\ &\quad + (g_{\mathrm{A}}^{\mathrm{e}})^2)] \cdot \eta_Z xF_3^Z \end{aligned} \tag{9.9.11}$$

The purpose for us to give the above result is only to show that, the structure function for polarized lepton scattering on polarized nucleon is more complicated, it is by no means as simple as unpolarized lepton–nucleon deep inelastic scattering. The details for the calculation and derivation can be referred to in the review article by PDG [55] and the references therein.

9.9.2. *The prediction of QCD to the variation of structure function with Q^2*

In the parton model, the structure function $F_i(x, Q^2) \to F_i(x)$ as $Q^2 \to \infty$, $\nu \to \infty$ and x being a fixed variable in the infinite momentum frame of the proton. Such scaling behavior is obtained by dropping the transverse momentum of the partons in the infinite momentum frame of the proton. But gluon can be radiated from quarks in QCD, which can violate the scaling behavior, and the violation is logarithmic in QCD. The radiated gluon can split into $q\bar{q}$ as Q^2 increasing, which makes the distribution of the initial quark momentum soften and the sea of $q\bar{q}$ increase. The smaller x and the larger Q^2 is, the stronger such process is. The variation of $F_2(x, Q^2)$ with Q^2 with $x \sim 0.25$ is shown in Fig. 9.18 [56]. The scaling behavior

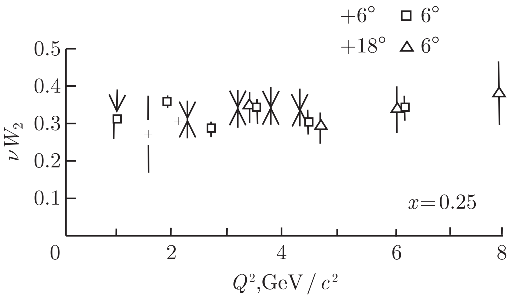


Fig. 9.18. The structure function of the proton $F_2(x, Q^2)$ varies with Q^2, which shows apparent scaling behavior.

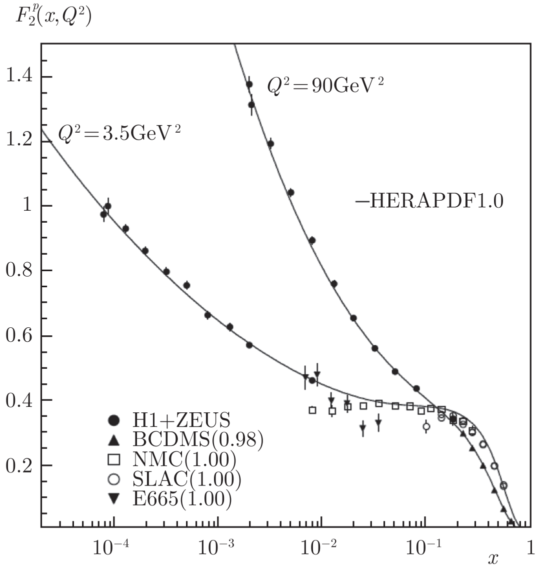


Fig. 9.19. The experimental curve for the structure function of the proton $F_2^p(x, Q^2)$ at $Q^2 = 3.5\,\text{GeV}^2$ and $Q^2 = 90\,\text{GeV}^2$, varying with x, where different data points come from different experimental groups. The solid curve is the parametrization form by HERAPDF group. It is shown that scaling behavior appears around $x \sim 0.14$.

is apparently remained, and $F_2(x, Q^2)$ is almost independent of Q^2. But when x goes smaller the situation is different. The experimental curve in Fig. 9.19 clearly shows this issue [55].

The process mentioned above can be described with the variation of the parton distribution function (PDF) $f_a(x, \mu^2)$, where a stands for q, $\bar{\text{q}}$ or gluon, and $\mu^2 = Q^2$ is the momentum transfer squared in the scattering (μ is the energy scale of the process). $f_a(x, \mu^2)$ is the distribution density of parton a in the proton with momentum xP and transverse momentum integrated from zero to the scale μ.

For $Q^2 \gg M^2$, the structure function F_i can be described with PDF

$$F_i = \sum_i C_i^a \otimes f_a \tag{9.9.12}$$

where $\otimes$ stands for convolution integral

$$C \otimes f = \int_x^1 \frac{\mathrm{d}y}{y} C(y) f\left(\frac{x}{y}\right) \tag{9.9.13}$$

with C_i^a being the expansion series over α_s. The variation of $f_a(x, \mu^2)$ with respect to μ^2 can be described by DGLAP equation

$$\frac{\partial f_a(x, \mu^2)}{\partial \ln \mu^2} \sim \frac{\alpha_s(\mu^2)}{2\pi} \sum_b P_{ab} \otimes f_b(x, \mu^2) \tag{9.9.14}$$

where P_{ab} stands for the probability of parton b fragmenting into a (the probability $b \to a$). The initial condition is needed when solving Eq. (9.9.14), i.e. the explicit form of $f_a(x, \mu_0^2)$ is needed. But this initial condition can not be calculated directly in QCD at present because of its nonperturbative characteristic. It can only be determined from experiment. If the initial PDF is known, the variation of PDF with $Q^2 = \mu^2$ can be predicted by Eq. (9.9.14) in QCD, and the value of PDF — $f_a(x, \mu^2)$ at any other scale μ can be obtained.

At present the parameterization forms of parton distribution function for proton and neutron have been given by six different groups [55], where the initial distribution, i.e. the form of $f_a(x, Q_0^2)$ at a scale $\mu^2 = Q_0^2$, was speculated at first, then $xf_a(x, Q^2)$ was calculated and compared with experimental data of processes $lp \to l'\mathrm{X}$ and $pp \to (W^\pm \to l^\pm \nu)\mathrm{X}$ etc. This makes the parametrization form more close to reality. Then this parametrization form of PDF can be used to predict other processes. The details can be found in [55] and the references therein, which will not be touched further here.

9.9.3. *Examples for application*

In this section we illustrate how the parton distribution function can be used in perturbation calculation in the framework of perturbation QCD.

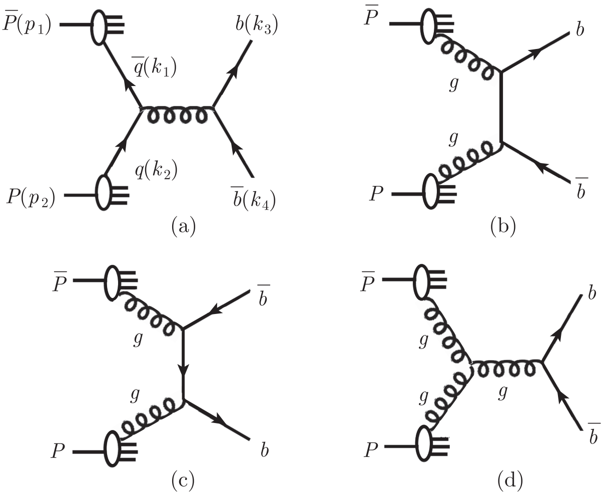

$\overline{P}(p_1)$
$\overline{q}(k_1)$
$b(k_3)$
$q(k_2)$
$P(p_2)$
$\overline{b}(k_4)$
(a)
$\overline{P}$
b
g
g
P
$\overline{b}$
(b)
$\overline{P}$
$\overline{b}$
g
g
P
b
(c)
$\overline{P}$
b
g
g
g
P
$\overline{b}$
(d)

Fig. 9.20. The lowest-order Feynman diagrams of the process of $p\bar{p} \to b\bar{b}$X.

We only discuss the process of $p\bar{p} \to b\bar{b}$X [57]. The Feynman diagrams are shown in Fig. 9.20.

As shown in Fig. 9.20(a), let $k_1 = xp_1$, $k_2 = yp_2$. The energy conservation for the subprocess gives

$$(xp_1 + yp_2)^2 = (k_3 + k_4)^2 \tag{9.9.15}$$

The center-of-mass energy squared of $\bar{p}p$ collision is $s = (p_1 + p_2)^2$. When s is larger than a few hundreds GeV^2, the left-hand side of Eq. (9.9.15) $\sim xys$, so the condition for generating $b\bar{b}$ pair is

$$xys \geq 4m_b^2 \tag{9.9.16}$$

The total cross section of $p\bar{p} \to b\bar{b}$X includes two parts, which are the cross section of $q\bar{q} \to b\bar{b}$, $\sigma_{\text{tot}}^{q\bar{q}\to b\bar{b}}$ and that of $gg \to b\bar{b}$, $\sigma_{\text{tot}}^{gg\to b\bar{b}}$

$$\begin{aligned}\sigma_{\text{tot}}^{q\bar{q}\to b\bar{b}}(p\bar{p} \to b\bar{b}\text{X}) &= \iint_{xys\geq 4m_b^2} \mathrm{d}x\mathrm{d}y\hat{\sigma}_{q\bar{q}\to b\bar{b}}\bar{q}^{\bar{p}}(x, Q^2)q^p(y, Q^2)\\ &= \iint_{xys\geq 4m_b^2} \mathrm{d}x\mathrm{d}y\hat{\sigma}_{q\bar{q}\to b\bar{b}}\frac{1}{xy}[x\bar{q}^{\bar{p}}(x, Q^2)\cdot yq^p(y, Q^2)]\end{aligned} \tag{9.9.17}$$

where $\bar{q}^{\bar{p}}(x, Q^2)$ is the number of the antiquarks with momentum xp_1 in the antiproton $\bar{p}$, and $q^p(y, Q^2)$ is the number of quarks with momentum yp_2 in the proton. We extract a factor $1/(xy)$ in front of Eq. (9.9.17) because the parameterization of the structure functions or parton distribution function PDF are always given in the form of $xf(x, Q^2)$. What given in the square bracket [ ] are just such parameterization form. The values of them can be given in six different parameterization forms [55], and can be used in practical calculations.

The contribution of Figs. 9.20(b)–(d) is

$$\begin{aligned}\sigma_{\mathrm{tot}}^{gg\to b\bar{b}}(p\bar{p} \to b\bar{b}\mathrm{X}) &= \iint_{xys\geq 4m_b^2} \mathrm{d}x\mathrm{d}y\hat{\sigma}_{gg\to b\bar{b}} G^{\bar{p}}(x, Q^2)G^p(y, Q^2) \\ &= \iint_{xys\geq 4m_b^2} \mathrm{d}x\mathrm{d}y\hat{\sigma}_{gg\to b\bar{b}} \frac{1}{xy}[xG^{\bar{p}}(x, Q^2)\cdot yG^p(y, Q^2)]\end{aligned} \tag{9.9.18}$$

where $G^{\bar{p}}(x, Q^2)$ is the distribution function of gluon with momentum xp_1 in the antiproton, and $G^p(y, Q^2)$ the distribution function of gluon with momentum yp_2 in proton.

The total cross sections of parton scattering $\hat{\sigma}_{q\bar{q}\to b\bar{b}}$ and $\hat{\sigma}_{gg\to b\bar{b}}$ in the above two equations can be calculated in perturbation theory [57]. What is needed to note is that the contribution of gluon scattering, i.e. $gg \to b\bar{b}$ process, can be neglected when the energy in the collision is small, at the same time, the contribution of sea quarks are not large either, so only the valence quarks' contribution needs to be considered, because both x and y are larger at this time ($x \sim y \sim 0.3$). What is discussed in [57] is just the same case. But, if the collision energy is very large, such as the energy of LHC for which $x, y \ll 1$, then the gluon scattering process $gg \to b\bar{b}$ dominates, only Eq. (9.9.18) is necessary to be calculated, which can be seen from Fig. 9.21 [58]. For different processes and different energies, which subprocess in Fig. 9.20 should be considered and which PDF is selected should be determined according to the specific situation. The method for selection is given in [57], to which the interested readers can refer to.

In Secs. 9.10 and 9.11 we will discuss some topics about nonperturbative QCD. We will first discuss bound state problem, then discuss a little about lattice gauge theory.

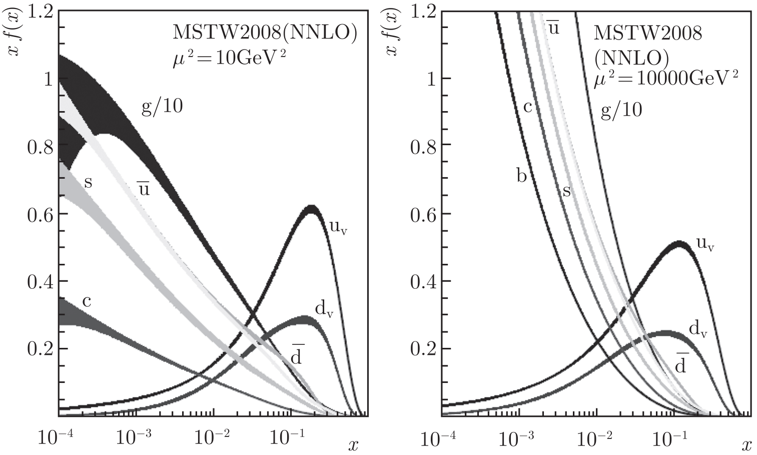


Fig. 9.21. The product of $xf(x)$ as a function of x. $f(x)$ is the unpolarized parton distribution function in proton. The distributions of valence quark, sea quark and gluon in proton as a function of x are given in the figure with $\mu^2 = 10\,\mathrm{GeV}^2$ and $10^4\,\mathrm{GeV}^2$.

9.10. A Brief Introduction to Some Bound States in QCD

9.10.1. e^+e^- *collision, charmonium and bottomonium*

The most important advantage of e^+e^- collider is to generate particles through electromagnetic interaction of e^+e^- with low background. There have been many constructed and in running e^+e^- colliders in the world, including energy symmetric e^+e^- collider, such as SPEAR at SLAC in the United States of America, CESR at Cornell University, BEPC in China and LEP in Europe, etc. There are also energy asymmetric e^+e^- collider, such as PEP-II in the United States of America, KEK-B in Japan, etc. For the details of e^+e^- colliders existing in the world, such as, name and colliding energy, etc., the readers can refer to the review article of PDG [59].

The energy level and decays of bound states of c–$\bar{c}$ (charmonium) and b–$\bar{b}$ (bottomonium) have been thoroughly measured in e^+e^- colliding experiments, which indicate that the energy levels of both charmonium and bottomonium can be described by non-relativistic quantum mechanics. For example, if the interaction potential of $c - \bar{c}$ is taken as [60]

$$V(r) = -\frac{k}{r} + \frac{r}{a^2} \tag{9.10.1}$$

the energy levels of charmonium can be well generated in theory. The first term in Eq. (9.10.1) is for Coulomb potential describing the short-distance interaction between c–$\bar{c}$, and the linear potential $\frac{r}{a^2}$ describes the long-distance interaction, which is required by the color confinement of QCD. Therefore, Eq. (9.10.1) is the phenomenological potential that is consistent with the characteristic of QCD.

The diagrams of the energy level of charmonium and bottomonium measured in experiment are given in Figs. 9.22 and 9.23, respectively [61].

The solid lines in Fig. 9.22 stand for the confirmed states in experiment, of which η_c and h_c are singlet states, while ψ and χ_{cJ} are triplet states. The not fully named charmonium-like states are denoted as X. What given in the parentheses such as (1S), (2P) etc. are the radial quantum number and the orbital angular momentum. The hadronic transitions are given in Fig. 9.22, while the single-photon transitions, such as $\psi(ns) \to \gamma\eta_c(mp)$, $\psi(ns) \to \gamma\chi_{cJ}(mp)$ and $\chi_{cJ}(1P) \to \gamma J/\psi$ are not labeled.

The solid lines in Fig. 9.23 stand for the confirmed states in experiment, and η_b and h_b are singlet states, while Υ and χ_{bJ} are triplet states. What given in parentheses are the radial quantum number and the orbital angular momentum. For example, $h_b(2P)$ denotes 2^1P_1 state with ($n = 2$, $L = 1$, $S = 0$, $J = 1$, $PC = +-$). The hadronic transitions are also labeled in Fig. 9.23, but the single-photon transitions, such as $\Upsilon(ns) \to \gamma\eta_b(mp)$, $\Upsilon(ns) \to \gamma\chi_{bJ}(mp)$ and $\chi_{bJ}(nP) \to \gamma\Upsilon(ms)$ etc. are not marked.

The theoretical and experimental study of the energy levels and decay properties of charmonium and bottomonium still continues. The interested readers may continue to follow the frontier developments in this field.

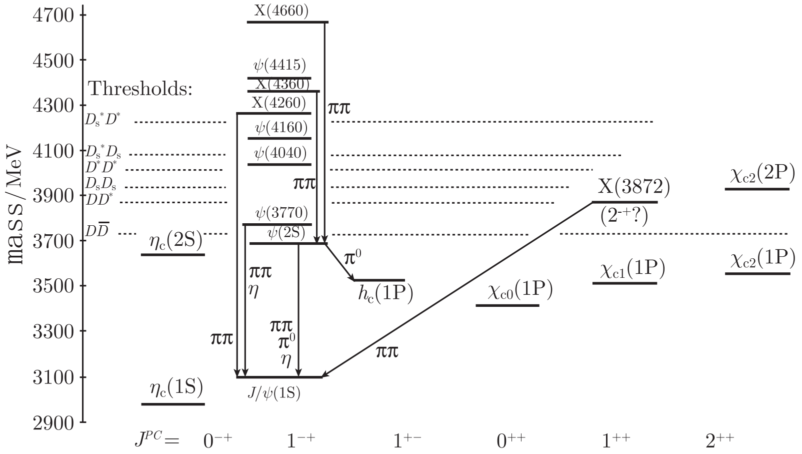

4700
4500
4300
4100
3900
3700
3500
3300
3100
2900
mass/MeV
Thresholds:
$D_s^*D^*$
$D_s^*D_s$
D^*D^*
D_sD_s
DD^*
$D\bar{D}$
X(4660)
ψ(4415)
X(4360)
X(4260)
ψ(4160)
ψ(4040)
ψ(3770)
ψ(2S)
J/ψ(1S)
η_c(2S)
η_c(1S)
ππ
ππ
π⁰
h_c(1P)
ππ
η
ππ
ππ
π⁰
η
ππ
X(3872)
(2⁻⁺?)
χ_{c0}(1P)
χ_{c1}(1P)
χ_{c2}(2P)
χ_{c2}(1P)
J^{PC}= 0⁻⁺ 1⁻⁺ 1⁺⁻ 0⁺⁺ 1⁺⁺ 2⁺⁺

Fig. 9.22. The energy level and decays of the charmonium.

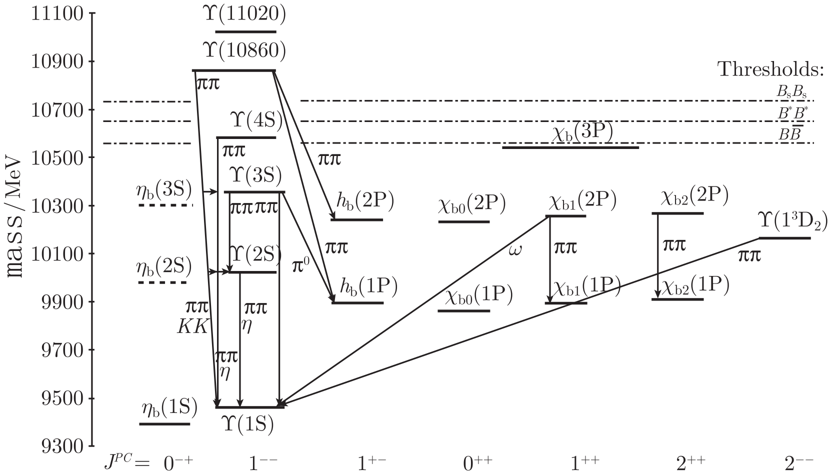

mass/MeV
11100
10900
10700
10500
10300
10100
9900
9700
9500
9300
Υ(11020)
Υ(10860)
Thresholds:
B_sB_s
B^*B^*
$B\bar{B}$
Υ(4S)
$\chi_b(3P)$
Υ(3S)
ππ
ππππ
$\eta_b(3S)$
$\eta_b(2S)$
Υ(2S)
π^0
$h_b(2P)$
$h_b(1P)$
$\chi_{b0}(2P)$
$\chi_{b0}(1P)$
$\chi_{b1}(2P)$
$\chi_{b1}(1P)$
$\chi_{b2}(2P)$
$\chi_{b2}(1P)$
$\Upsilon(1^3D_2)$
ω
KK
η
$\eta_b(1S)$
Υ(1S)
$J^{PC} = 0^{-+}$
1^{--}
1^{+-}
0^{++}
1^{++}
2^{++}
2^{--}

Fig. 9.23. The energy level and decays of the bottomonium.

9.10.2. *Glueball and hybrid state*

It can be seen from the Lagrangian of QCD given previously that, although gluon is massless, there is strong interaction between gluons which is different from properties of photon, so gluons should have bound state, that is, a glueball composed of pure gluons. A glueball can be bound state of not only two gluons, but also three or more gluons (gg, $ggg, \ldots$). Besides glueball, according to QCD, there should also be hybrid states, such as $q\bar{q}g$, $q\bar{q}gg$, etc. One of the most crucial tests for the low-energy properties of QCD is whether the existence of glueball and hybrid states can be confirmed experimentally. The C parity of glueball and hybrid states can be discussed in the following manner [62]. The Lagrangian of QCD is invariant under transformation of charge conjugation ($\hat{C}$ transformation). One term in QCD is $\bar{\psi}\gamma^\mu\left(\frac{\lambda^a}{2}\right)A^a_\mu\psi$, the $\hat{C}$ transformation of which is

$$\hat{C}\bar{\psi}\gamma^\mu\frac{\lambda^a}{2}A^a_\mu\psi\hat{C}^{-1} = -\bar{\psi}\gamma^\mu\frac{\lambda^{a*}}{2}\bar{A}^a_\mu\psi \tag{9.10.2}$$

where

$$\bar{A}^a_\mu = \hat{C}A^a_\mu\hat{C}^{-1} \tag{9.10.3}$$

$\hat{C}$ transformation includes transformations in Dirac spinor space and color space.

It can be apparently known from Eqs. (9.10.2) and (9.10.3) that $\hat{C}$ invariance requires

$$\lambda^{a*}\bar{A}_\mu^a = -\lambda^a A_\mu^a \tag{9.10.4}$$

However

$$\begin{cases} \lambda^{a*} = -\lambda^a & a = 2,5,7 \\ \lambda^{a*} = \lambda^a & a = 1,3,4,6,8 \end{cases} \tag{9.10.5}$$

so Eq. (9.10.4) gives

$$\begin{cases} \bar{A}_\mu^a = \delta(a)A_\mu^a \\ \delta(a) = \begin{cases} 1 & \text{for } a = 2,5,7 \\ -1 & \text{for } a = 1,3,4,6,8 \end{cases} \end{cases} \tag{9.10.6}$$

Since under $\hat{C}$ transformation $\bar{\psi}\gamma_\mu\psi \to -\bar{\psi}\gamma_\mu\psi$, Eq. (9.10.2) can be equivalently viewed as

$$\lambda^a A_\mu^a \xrightarrow{\hat{C}} -\lambda^a A_\mu^a \tag{9.10.7}$$

under $\hat{C}$ transformation, which is similar to the $\hat{C}$ transformation of photon field A_μ

$$A_\mu \xrightarrow{\hat{C}} -A_\mu \tag{9.10.8}$$

The operator of glueball mast be written in a way of $\mathrm{SU}(3)_C$ gauge invariance. For the state of two gluons gg, there are scalar operator

$$G_{\mu\nu}^a G^{a\mu\nu} \tag{9.10.9}$$

and pseudoscalar operator

$$G_{\mu\nu}^a \tilde{G}^{a\mu\nu} \tag{9.10.10}$$

where $G_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + f^{abc}A_\mu^b A_\nu^c$, $\tilde{G}^{a\mu\nu} = \varepsilon^{\mu\nu\rho\sigma}G_{\rho\sigma}^a$. Therefore, there are also three-gluon and four-gluon states. It can be seen from Eq. (9.10.6) that be

$$
\begin{gathered}
G^a_{\mu\nu} \xrightarrow{\hat{C}} \bar{G}^a_{\mu\nu} = \delta(a) G^a_{\mu\nu} \\
\begin{cases}
G^a_{\mu\nu} G^{a\mu\nu} \xrightarrow{\hat{C}} [\delta(a)]^2 \, G^a_{\mu\nu} G^{a\mu\nu} \\
G^a_{\mu\nu} \tilde{G}^{a\mu\nu} \xrightarrow{\hat{C}} [\delta(a)]^2 \, G^a_{\mu\nu} \tilde{G}^{a\mu\nu}
\end{cases}
\end{gathered}
\tag{9.10.11}
$$

That is, the $\hat{C}$ parities of the operators in Eqs. (9.10.9) and (9.10.10) are positive, so the quantum state of gg glueball with the lowest angular momentum can be 0^{++} state (see Eq. (9.10.9)) or 0^{-+} state (see Eq. (9.10.10)).

The two terms of three-gluon operators without derivative and being gauge invariant are (the Lorentz indices omitted)

$$f^{abc} A^a A^b A^c \tag{9.10.12}$$

$$d^{abc} A^a A^b A^c \tag{9.10.13}$$

where f^{abc} and d^{abc} are the symmetric and antisymmetric structure constants of $\mathrm{SU}(3)_C$ group, respectively. It can be known from Eq. (9.10.6) and the non-zero values of f^{abc} and d^{abc} that the $\hat{C}$ parity of $f^{abc}A^aA^bA^c$ is positive, and that of $d^{abc}A^aA^bA^c$ negative.

For a hybrid of $\bar{q}qg$, the relevant gauge invariant operator is

$$J_\mu = \bar{\psi}\gamma^\nu G^a_{\mu\nu} \frac{\lambda^a}{2} \psi \tag{9.10.14}$$

From Eqs. (9.10.12) and (9.10.7), it can be known that the $\hat{C}$ parity of Eq. (9.10.14) is positive. Under the transformation of parity $\hat{P}$, Eq. (9.10.14) transforms as a vector

$$\bar{\psi}\gamma^\nu G^a_{\mu\nu} \frac{\lambda^a}{2} \psi \xrightarrow{\hat{P}} \bar{\psi}\gamma_\nu G^{a\mu\nu} \frac{\lambda^a}{2} \psi \tag{9.10.15}$$

so, for Eq. (9.10.15) as $\mu = 1, 2, 3$, its $J^{PC} = 1^{-+}$, and as $\mu = 0$, $J^{PC} = 0^{++}$.

If using $\gamma^\nu\gamma^5$ to replace γ^ν in Eq. (9.10.15), i.e.

$$\bar{\psi}\gamma^\nu\gamma^5 G^a_{\mu\nu} \frac{\lambda^a}{2} \psi \tag{9.10.16}$$

which gives two hybrid states with $J^{PC} = 1^{+-}$ and 0^{-+}, we can also construct tensor operator, so that describing the hybrid of $J = 2$, which is not discussed further here.

Equation (9.10.15) includes the state of $J^{PC} = 1^{-+}$, which cannot be composed of $\bar{q}q$, so it is an exotic state.

There are many methods to calculate the mass spectra of glueball and hybrid in the literature, such as, the bag model, flux tube model, lattice QCD and QCD Sum Rule, etc. Next we shall present some theoretical predictions about these mass spectra.

For the mass of the lightest glueball, the state of 0^{++}, the predictions given by lattice QCD in quench approximation recently are $m_{++} = 1750(50)(80)\,\text{MeV}$ [63], $1710 \pm 100\,\text{MeV}$ [64], $1650\,\text{MeV}$ [65] and $1550\,\text{MeV}$ [66], etc.

For hybrid state of 1^{-+}, flux tube model predicts [67] $m_{1^{-+}} \approx 1.9\,\text{GeV}$, lattice QCD [68] $m_{1^{-+}} \sim 1.8 - 1.9\,\text{GeV}$, and Bag model [69] $m_{1^{-+}} \geq 1.4\,\text{GeV}$.

In experiment, glueball can mix with the $q\bar{q}$ state with the same quantum numbers and similar masses, so pure glueball does not exist in general, unless there is no $q\bar{q}$ state of similar mass near some glueball mass. For example, the meson of 0^{++} state with mass around 1500 MeV can mix with the pure ground state glueball of 0^{++} state whose mass is also near the same region, which can generate the experimentally observed mesons $f_0(1370)$, $f_0(1500)$ and $f_0(1710)$ [70].

It is speculated in experiment [71] that $\pi_1(1400)$ and $\pi_1(1600)$ can be the hybrid state of $J^{PC} = 1^{-+}$. But this is not confirmed. They may also be tetraquark states.

9.11. Nonperturbative QCD and Lattice Gauge Theory

What discussed in previous sections are mainly perturbative QCD or the high-energy behavior of QCD. Large amount of experiments indicate that the calculation of perturbative QCD is successful for process of large momentum transfer. The low-energy behavior of QCD or nonperturbative QCD has not been discussed yet. The coupling constant of strong interaction $\alpha_s \geq 1$ for low-energy process, such that perturbative calculation cannot apply. The most powerful non-perturbative theory is the lattice gauge theory, or lattice QCD (LQCD), developed in several decades. LQCD is defined on a hypercubic in a discretized 4-dimensional Euclidean space–time. The hypercubic can be expressed by $a(n_1, n_2, n_3, n_4)$, where a is the lattice spacing, n_1, n_2, n_3 are the lattice numbers in the 3-dimensional space, and n_4 the time lattice number. We can imagine that the Minkowski $3+1$-dimensional space–time can be transformed into 4-dimensional Euclidean space by rotating the time axis to a real axis with Wick rotation. It can be seen by comparing with the field theory defined in an infinite volume that, the smaller a (i.e. the denser the lattice and the larger n_i $(i=1,2,3,4)$), the closer the calculation to the limit of the continuous field theory. Since space–time is divided into lattices of finite length, ultraviolet divergence does not occur in the theory. This is why the lattice gauge theory can handle non-perturbative calculations.

In lattice QCD fermion fields are placed on lattice sites and gauge fields (gluon fields) on the links between cites. Lattice QCD should be kept gauge invariant, which can heavily restrict the operator form and greatly simplify the calculations.

Physical quantities are calculated with path integral method in LQCD. Firstly the action should be selected, then the Euclidean-space partition function is defined, finally physical quantities can be calculated with Monte-Carlo method or using two or three-point Green functions. Since the operators defined in LQCD depend on the quark–antiquark field, the fermion determinant shall be encountered when doing the path integral. The contribution of sea quarks was usually dropped in earlier calculation, which is called quench approximation. Thanks to advances in computers and computing technology, the contribution of sea quarks can now be considered.

At present the spacing of the hypercubic has been reached to $a(64^3\times 144)$, i.e. $a\sim(0.1\text{–}0.05)\,\mathrm{fm}$, $n_1=n_2=n_3=64$, $n_4=144$.

Currently, the following four kinds of physical quantities can be calculated in LQCD.

(1) The hadronic mass spectra: the masses of hadrons composed of $\mathrm{u,d,s}$ quarks, the masses of hadrons composed of heavy–light and heavy–heavy quark–antiquarks, and the masses of glueballs and hybrids.
(2) The decay constants f_π, f_K, f_D, f_{D_s}, f_B, f_{B_s} etc., and the bag parameters B_K, B_B, etc.
(3) The transition form factors, such as the form factors of the decays $K\to\pi l\nu$, $D\to Kl\nu$, $B\to\pi l\nu$, $B\to D^{(*)}l\nu$, etc.
(4) The strong coupling constant $\alpha_s(Q^2)$.

As computing power improves over the next few years, transition matrix elements, such as $\frac{\varepsilon'}{\varepsilon}$, may also be calculated in LQCD.

The errors in LQCD include two types: one is the statistical error, which is generated by Monte-Carlo numerical calculation and the sample size; the other is the systematic error, such as those resulted from the transition to the continuous limit, i.e. the error caused by the finite cubic volume. The quark masses of u, d, s used in the calculation of LQCD are generally large, the errors can also occur when the quark masses are extended to smaller current quark mass. Generally speaking, the above errors are basically controllable. The total error of LQCD is $\leq 6\%$, and the errors of some physical quantities are only about 2%. Therefore, the result of lattice QCD calculation poses a strong restriction on the measurable physical quantities.

For the detailed discussion about the above content, the readers can refer to the review article *Lattice Quantum Chromodynamics* in *Review of Particle Physics* by PDG group [72] and the references therein.

9.12. The Strong *CP* Problem

The Lagrangian of strong interaction given in Eq. (9.1.24) is $\mathrm{SU}(3)_\mathrm{C}$ gauge invariant, and it is also CP invariant simultaneously. However, gauge symmetry can not necessarily guarantee CP symmetry. Under the condition of respecting gauge symmetry, it is still possible that the following term exists in QCD

$$\theta \frac{\alpha_s}{8\pi} F^a_{\mu\nu} \tilde{F}^{a\mu\nu} \tag{9.12.1}$$

where

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f^{abc} A^b_\mu A^c_\nu \tag{9.12.2}$$

$$\tilde{F}^a_{\mu\nu} = \frac{1}{2} \varepsilon_{\mu\nu\rho\sigma} F^{a\rho\sigma} \tag{9.12.3}$$

and $\tilde{F}^a_{\mu\nu}$ is the dual tensor of $F^a_{\mu\nu}$.

We can verify that Eq. (9.12.1) is gauge invariant using the method given in Eqs. (9.1.20)–(9.1.22). Next we shall prove that Eq. (9.12.1) is not CP invariant. For this purpose, we need first to find out the transformation of gluon field A^a_μ under the parity and charge conjugate transformation. This can be done by the requirement of quark and gluon interaction in QCD being invariant under parity and charge conjugate transformation.

The interaction term of quark and gluon fields in QCD is

$$\mathcal{L}_I = \bar{\psi} \gamma^\mu T^a A^a_\mu \psi \tag{9.12.4}$$

Explicitly write the color indices α and β, then $\mathcal{L}_I$ can be written as

$$\mathcal{L}_I = \bar{\psi}_\alpha \gamma^\mu T^a_{\alpha\beta} \psi_\beta A^a_\mu \tag{9.12.5}$$

Using the parity transformation of fermion field derived in Sec. 2.3 of Chapter 2 (see Eq. (2.3.10)), i.e.

$$\hat{P}\psi(x)\hat{P}^{-1} = \mathrm{e}^{\mathrm{i}\phi_p}\gamma^0\psi(\tilde{x}) \tag{9.12.6}$$

$$\hat{P}\bar{\psi}(x)\hat{P}^{-1} = \mathrm{e}^{-\mathrm{i}\phi_p}\bar{\psi}(\tilde{x})\gamma^0 \tag{9.12.7}$$

we can get the parity transformation of the quark–gluon interaction term

$$\begin{aligned}\hat{P}\int \mathrm{d}^4x\mathcal{L}_I(x)\hat{P}^{-1} &= \int \mathrm{d}^4x\hat{P}\mathcal{L}_I(x)\hat{P}^{-1}\\ &= \int \mathrm{d}^4x\hat{P}\bar{\psi}(x)\gamma^\mu T^a\psi(x)A^a_\mu(x)\hat{P}^{-1}\\ &= \int \mathrm{d}^4x\hat{P}\bar{\psi}(x)\hat{P}^{-1}\gamma^\mu T^a\hat{P}\psi(x)\hat{P}^{-1}\hat{P}A^a_\mu(x)\hat{P}^{-1}\\ &= \int \mathrm{d}^4x\bar{\psi}(\tilde{x})\gamma^0\gamma^\mu T^a\gamma^0\psi(\tilde{x})\hat{P}A^a_\mu(x)\hat{P}^{-1}\\ &= \int \mathrm{d}^4x\bar{\psi}(\tilde{x})\gamma_\mu T^a\psi(\tilde{x})\hat{P}A^a_\mu(x)\hat{P}^{-1}\end{aligned} \tag{9.12.8}$$

In the last step of the above equation, use has been made of

$$\gamma^0\gamma^\mu\gamma^0 = \gamma_\mu \tag{9.12.9}$$

Equation (9.12.8) indicates that, as long as the parity transformation of gluon field is in the following way

$$\hat{P}A^a_\mu(x)\hat{P}^{-1} = A^{a\mu}(\tilde{x}) \tag{9.12.10}$$

the quark–gluon interaction term can be guaranteed to be invariant under parity transformation, i.e.

$$\hat{P}\int \mathrm{d}^4x\mathcal{L}_I(x)\hat{P}^{-1} = \int \mathrm{d}^4x\mathcal{L}_I(x) \tag{9.12.11}$$

Therefore, Eq. (9.12.10) is just the parity transformation of the gluon field, which is similar to the parity transformation of electromagnetic field.

Next let us derive the charge conjugate transformation of the gluon field.

Using the charge conjugate transformation of fermion field given in Sec. 2.4 of Chapter 2 (see Eq. (2.4.13)), i.e.

$$\begin{aligned}\hat{C}\psi(x)\hat{C}^{-1} &= \mathrm{e}^{\mathrm{i}\theta_C} C\bar{\psi}^T(x)\\ \hat{C}\bar{\psi}(x)\hat{C}^{-1} &= -\mathrm{e}^{-\mathrm{i}\theta_C}\psi^T(x)C^{-1}\end{aligned} \tag{9.12.12}$$

and the equation for the charge conjugation matrix

$$C^{-1}\gamma^\mu C = -\gamma^{\mu T} \tag{9.12.13}$$

we can find the charge conjugate transformation of $\mathcal{L}_I$ in Eq. (9.12.5)

$$\begin{aligned}\hat{C}\mathcal{L}_I\hat{C}^{-1} &= \hat{C}\bar{\psi}_\alpha\gamma^\mu T^a_{\alpha\beta}\psi_\beta A^a_\mu\hat{C}^{-1}\\ &= \hat{C}\bar{\psi}_\alpha\hat{C}^{-1}\gamma^\mu T^a_{\alpha\beta}\hat{C}\psi_\beta\hat{C}^{-1}\hat{C}A^a_\mu\hat{C}^{-1}\\ &= -\psi^T_\alpha C^{-1}\gamma^\mu T^a_{\alpha\beta}C\bar{\psi}^T_\beta\hat{C}A^a_\mu\hat{C}^{-1}\\ &= -T^a_{\alpha\beta}\psi^T_\alpha C^{-1}\gamma^\mu C\bar{\psi}^T_\beta\hat{C}A^a_\mu\hat{C}^{-1}\\ &= T^a_{\alpha\beta}\psi^T_\alpha\gamma^{\mu T}\bar{\psi}^T_\beta\hat{C}A^a_\mu\hat{C}^{-1}\\ &= T^a_{\alpha\beta}(\bar{\psi}_\beta\gamma^\mu\psi_\alpha)^T\hat{C}A^a_\mu\hat{C}^{-1}\\ &= -T^a_{\alpha\beta}\bar{\psi}_\beta\gamma^\mu\psi_\alpha\hat{C}A^a_\mu\hat{C}^{-1}\\ &= -(T^{aT})_{\beta\alpha}\bar{\psi}_\beta\gamma^\mu\psi_\alpha\hat{C}A^a_\mu\hat{C}^{-1}\\ &= -\bar{\psi}\gamma^\mu T^{aT}\psi\hat{C}A^a_\mu\hat{C}^{-1}\end{aligned} \tag{9.12.14}$$

That the term $\bar{\psi}_\beta\gamma^\mu\psi_\alpha$ is a C-number has been considered in deriving the above equation and the transposition does not change its value, but there should be $(\bar{\psi}_\beta\gamma^\mu\psi_\alpha)^T = -\bar{\psi}_\beta\gamma^\mu\psi_\alpha$ because of the anti-commutativity of fermions.

It can be seen from the Gell-Mann matrix given in Eq. (5.1.16) that the generators of SU(3) group satisfy the following relation:

$$\begin{cases} T^{aT} = -T^a, & (a = 2, 5, 7)\\ T^{aT} = T^a, & (a = 1, 3, 4, 6, 8)\end{cases} \tag{9.12.15}$$

The above equation can also be written as

$$\begin{cases} T^{aT} = -\delta(a)T^a\\ \delta(a) = \begin{cases} 1, & (a = 2, 5, 7)\\ -1, & (a = 1, 3, 4, 6, 8)\end{cases}\end{cases} \tag{9.12.16}$$

Substituting the result of Eq. (9.12.16) into Eq. (9.12.14), we can obtain

$$\hat{C}\mathcal{L}_I\hat{C}^{-1} = \bar{\psi}\gamma^\mu T^a \psi \delta(a)\hat{C}A_\mu^a\hat{C}^{-1} \tag{9.12.17}$$

Therefore, to keep the quark–gluon interaction invariant under charge conjugate transformation, there must be

$$\hat{C}A_\mu^a\hat{C}^{-1} = \delta(a)A_\mu^a \tag{9.12.18}$$

This is the charge conjugate transformation of the gluon field, which is similar to the charge conjugate transformation of a photon, except that it differs by a coefficient $\delta(a)$ for different gluons. Equation (9.12.18) is the same as Eq. (9.10.6). Here a more detailed derivation is given.

With Eqs. (9.12.10) and (9.12.18) in hand, we can analyze the CP transformation of the pure gluon terms $F_{\mu\nu}^a F^{a\mu\nu}$ and $F_{\mu\nu}^a \tilde{F}^{a\mu\nu}$. Let us analyze the term $F_{\mu\nu}^a F^{a\mu\nu}$ at first. Take parity transformation of the term $\int \mathrm{d}^4 x F_{\mu\nu}^a(x) F^{a\mu\nu}(x)$ in the action

$$\begin{aligned}\hat{P}\int \mathrm{d}^4 x F_{\mu\nu}^a(x)F^{a\mu\nu}(x)\hat{P}^{-1} &= \int \mathrm{d}^4 x \hat{P}F_{\mu\nu}^a(x)\hat{P}^{-1}\hat{P}F^{a\mu\nu}(x)\hat{P}^{-1}\\ &= \int \mathrm{d}^4 x F^{a\mu\nu}(\tilde{x})F_{\mu\nu}^a(\tilde{x})\\ &= \int \mathrm{d}^4 x F^{a\mu\nu}(x)F_{\mu\nu}^a(x)\end{aligned} \tag{9.12.19}$$

This indicates that the parity of $F_{\mu\nu}^a F^{a\mu\nu}$ is positive. Next we are going to find the charge conjugation of this term in the following.

We can directly act the charge conjugation on the term $F_{\mu\nu}^a F^{a\mu\nu}$ since charge conjugation has nothing to with space–time coordinate

$$\begin{aligned}\hat{C}F_{\mu\nu}^a F^{a\mu\nu}\hat{C}^{-1} &= \hat{C}F_{\mu\nu}^a\hat{C}^{-1}\hat{C}F^{a\mu\nu}\hat{C}^{-1}\\ &= \delta(a)^2 F_{\mu\nu}^a F^{a\mu\nu}\\ &= F_{\mu\nu}^a F^{a\mu\nu}\end{aligned} \tag{9.12.20}$$

Use has been made of

$$\hat{C}F_{\mu\nu}^a\hat{C}^{-1} = \delta(a)F_{\mu\nu}^a \tag{9.12.21}$$

in Eq. (9.12.20). To obtain this relation, the specific non-zero values of the structure constants of SU(3) group need to be considered in deriving Eq. (9.12.21), which can be found in Table 5.1.

Equations (9.12.19) and (9.12.20) show that $F^a_{\mu\nu}F^{a\mu\nu}$ in QCD is both P and C invariant, so it is certainly CP invariant.

And then let us think about P and C transformation for $F^a_{\mu\nu}\tilde{F}^{a\mu\nu}$ term. We can first prove the action term $\int \mathrm{d}^4xF^a_{\mu\nu}(x)\tilde{F}^{a\mu\nu}(x)$ is a pseudoscalar.

Proof. Summation is taken over all the possible values of the repeated indices μ, ν in $F^a_{\mu\nu}\tilde{F}^{a\mu\nu}$. But, because of the existence of the antisymmetric tensor $\varepsilon_{\mu\nu\rho\sigma}$ in $\tilde{F}^{a\mu\nu}$, only the unequal values of μ and ν are needed

$$(\mu,\nu) = (0,i),(i,0),(i,j)$$

where $(i,j) = 1,2,3$, and $i \neq j$. Therefore

$$F^a_{\mu\nu}\tilde{F}^{a\mu\nu} = F^a_{0i}\tilde{F}^{a0i} + F^a_{i0}\tilde{F}^{ai0} + F^a_{ij}\tilde{F}^{aij} \tag{9.12.22}$$

Take the parity transformation of the relevant action term

$$\begin{aligned}
&\hat{P}\int \mathrm{d}^4xF^a_{\mu\nu}(x)\tilde{F}^{a\mu\nu}(x)\hat{P}^{-1}\\
&\quad= \int \mathrm{d}^4x\hat{P}F^a_{\mu\nu}(x)\tilde{F}^{a\mu\nu}(x)\hat{P}^{-1}\\
&\quad= \int \mathrm{d}^4x\hat{P}F^a_{0i}(x)\tilde{F}^{a0i}(x)\hat{P}^{-1} + \int \mathrm{d}^4x\hat{P}F^a_{i0}(x)\tilde{F}^{ai0}(x)\hat{P}^{-1}\\
&\qquad+ \int \mathrm{d}^4x\hat{P}F^a_{ij}(x)\tilde{F}^{aij}(x)\hat{P}^{-1}
\end{aligned} \tag{9.12.23}$$

The first term in the above is

$$\begin{aligned}
\int \mathrm{d}^4x\hat{P}F^a_{0i}(x)\tilde{F}^{a0i}(x)\hat{P}^{-1} &= \int \mathrm{d}^4x\hat{P}F^a_{0i}(x)\hat{P}^{-1}\hat{P}\tilde{F}^{a0i}(x)\hat{P}^{-1}\\
&= \frac{1}{2}\int \mathrm{d}^4x\hat{P}F^a_{0i}(x)\hat{P}^{-1}\varepsilon^{0i\rho\sigma}\hat{P}F^a_{\rho\sigma}(x)\hat{P}^{-1}\\
&= \frac{1}{2}\int \mathrm{d}^4x\hat{P}F^a_{0i}(x)\hat{P}^{-1}\varepsilon^{0ijk}\hat{P}F^a_{jk}(x)\hat{P}^{-1}\\
&= \frac{1}{2}\int \mathrm{d}^4x(-F^a_{0i}(\tilde{x}))\varepsilon^{0ijk}F^a_{jk}(\tilde{x})\\
&= \frac{1}{2}\int \mathrm{d}^4x(-F^a_{0i}(x))\varepsilon^{0ijk}F^a_{jk}(x)\\
&= -\int \mathrm{d}^4xF^a_{0i}(x)\tilde{F}^{a0i}(x)
\end{aligned} \tag{9.12.24}$$

where $k = 1,2,3$. □

The second term in Eq. (9.12.23) is equal to the first term, so

$$\int \mathrm{d}^4x\hat{P}F^a_{i0}(x)\tilde{F}^{ai0}(x)\hat{P}^{-1} = -\int \mathrm{d}^4xF^a_{i0}(x)\tilde{F}^{ai0}(x) \tag{9.12.25}$$

The third term of Eq. (9.12.23) is

$$\begin{aligned}
&\int \mathrm{d}^4x\hat{P}F^a_{ij}(x)\tilde{F}^{aij}(x)\hat{P}^{-1}\\
&\quad= \int \mathrm{d}^4x\hat{P}F^a_{ij}(x)\hat{P}^{-1}\hat{P}\tilde{F}^{aij}(x)\hat{P}^{-1}\\
&\quad= \frac{1}{2}\int \mathrm{d}^4x\hat{P}F^a_{ij}(x)\hat{P}^{-1}\hat{P}\varepsilon^{ij\rho\sigma}F^a_{\rho\sigma}(x)\hat{P}^{-1}\\
&\quad= \frac{1}{2}\int \mathrm{d}^4x\hat{P}F^a_{ij}(x)\hat{P}^{-1}\hat{P}(\varepsilon^{ij0k}F^a_{0k}(x)+\varepsilon^{ijk0}F^a_{k0}(x))\hat{P}^{-1}\\
&\quad= \frac{1}{2}\int \mathrm{d}^4x\hat{P}F^a_{ij}(x)\hat{P}^{-1}(\varepsilon^{ij0k}\hat{P}F^a_{0k}(x)\hat{P}^{-1}+\varepsilon^{ijk0}\hat{P}F^a_{k0}(x)\hat{P}^{-1})\\
&\quad= \frac{1}{2}\int \mathrm{d}^4xF^a_{ij}(\tilde{x})(-\varepsilon^{ij0k}F^a_{0k}(\tilde{x})-\varepsilon^{ijk0}F^a_{k0}(\tilde{x}))\\
&\quad= \frac{1}{2}\int \mathrm{d}^4xF^a_{ij}(x)(-\varepsilon^{ij0k}F^a_{0k}(x)-\varepsilon^{ijk0}F^a_{k0}(x))\\
&\quad= \frac{1}{2}\int \mathrm{d}^4xF^a_{ij}(x)(-\varepsilon^{ij\rho\sigma}F^a_{\rho\sigma}(x) = -\int \mathrm{d}^4xF^a_{ij}(x)\tilde{F}^{aij}(x)
\end{aligned} \tag{9.12.26}$$

Substituting the results in Eqs. (9.12.24)–(9.12.26) back into Eq. (9.12.23), we can obtain

$$\hat{P}\int \mathrm{d}^4xF^a_{\mu\nu}(x)\tilde{F}^{a\mu\nu}(x)\hat{P}^{-1} = -\int \mathrm{d}^4xF^a_{\mu\nu}(x)\tilde{F}^{a\mu\nu}(x) \tag{9.12.27}$$

This indicates that $\int \mathrm{d}^4xF^a_{\mu\nu}(x)\tilde{F}^{a\mu\nu}(x)$ is indeed a pseudoscalar term.

Next let us consider the C transformation of $F^a_{\mu\nu}\tilde{F}^{a\mu\nu}$

$$\hat{C}F^a_{\mu\nu}\tilde{F}^{a\mu\nu}\hat{C}^{-1} = \delta(a)^2F^a_{\mu\nu}\tilde{F}^{a\mu\nu} = F^a_{\mu\nu}(x)\tilde{F}^{a\mu\nu}(x) \tag{9.12.28}$$

Jointly considering Eqs. (9.12.27) and (9.12.28), we can know that, under CP transformation, there is

$$\int \mathrm{d}^4xF^a_{\mu\nu}(x)\tilde{F}^{a\mu\nu}(x) \xrightarrow{CP} -\int \mathrm{d}^4xF^a_{\mu\nu}(x)\tilde{F}^{a\mu\nu}(x) \tag{9.12.29}$$

This shows that $F^a_{\mu\nu}\tilde{F}^{a\mu\nu}$ is not only P but also CP violating, but it is C invariant.

For the sake of the possible existence of the CP violating term, the Lagrangian of QCD can be extended to be

$$\mathcal{L}_{\mathrm{QCD}} = -\tfrac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} + \sum_{j=1}^{n}\left(\bar{\psi}_j i\gamma^\mu D_\mu\psi_j - m_j\bar{\psi}_{\mathrm{L}j}\psi_{\mathrm{R}j} - m_j^*\bar{\psi}_{\mathrm{R}j}\psi_{\mathrm{L}j}\right) + \theta\tfrac{\alpha_s}{8\pi}F^a_{\mu\nu}\tilde{F}^{a\mu\nu} \tag{9.12.30}$$

where ψ_j stands for quark field, and the left and right-handed fields for quarks are defined as

$$\psi_{\mathrm{L}j} = \frac{1}{2}(1-\gamma_5)\psi_j, \quad \psi_{\mathrm{R}j} = \frac{1}{2}(1+\gamma_5)\psi_j \tag{9.12.31}$$

As the quark masses are real $m_j^* = m_j$, the mass term of Eq. (9.12.30) goes back to the form of $m_j\bar{\psi}_j\psi_j$. We do not take $m_j^* = m_j$ for Eq. (9.12.30) at present, which is equivalent to keep the possibility that the quark mass may have a non-zero complex phase in some cases.

As the quark mass $m_j \to 0$, the Lagrangian of QCD is invariant under the following transformation

$$\psi_j \to \mathrm{e}^{\mathrm{i}\alpha_j\gamma_5}\psi_j, \quad j = 1, \ldots, n \tag{9.12.32}$$

This is called chiral symmetry, and the transformation of Eq. (9.12.32) is named as chiral transformation. As the quark mass is non-zero, the existence of the mass term violates the chiral symmetry. Now let us derive how the quark mass term changes under chiral transformation. For this purpose, we first expand the transformation factor $\mathrm{e}^{\mathrm{i}\alpha_j\gamma_5}$ to get another expression for it

$$\begin{aligned}
\mathrm{e}^{\mathrm{i}\alpha_j\gamma_5} &= \sum_{n=0}^{\infty}\frac{1}{n!}(\mathrm{i}\alpha_j\gamma_5)^n \\
&= \sum_{m=0}^{\infty}\frac{1}{(2m)!}(\mathrm{i}\alpha_j\gamma_5)^{2m} + \sum_{m=0}^{\infty}\frac{1}{(2m+1)!}(\mathrm{i}\alpha_j\gamma_5)^{2m+1} \\
&= \sum_{m=0}^{\infty}\frac{1}{(2m)!}(-1)^m\alpha_j^{2m} + \mathrm{i}\gamma_5\sum_{m=0}^{\infty}\frac{1}{(2m+1)!}(-1)^m\alpha_j^{2m+1} \\
&= \cos\alpha_j + \mathrm{i}\gamma_5\sin\alpha_j
\end{aligned} \tag{9.12.33}$$

i.e.

$$\mathrm{e}^{\mathrm{i}\alpha_j\gamma_5} = \cos\alpha_j + \mathrm{i}\gamma_5\sin\alpha_j \tag{9.12.34}$$

Using this equation, we can easily obtain the chiral transformation of $\psi_{\mathrm{L}j}$ and $\psi_{\mathrm{R}j}$

$$\begin{aligned}\psi_{\mathrm{L}j} &\to \frac{1}{2}(1-\gamma_5)\mathrm{e}^{\mathrm{i}\alpha_j\gamma_5}\psi_j \\ &= \frac{1}{2}(1-\gamma_5)(\cos\alpha_j + \mathrm{i}\gamma_5\sin\alpha_j)\psi_j \\ &= \left[\frac{1}{2}(1-\gamma_5)\cos\alpha_j - \frac{\mathrm{i}}{2}(1-\gamma_5)\sin\alpha_j\right]\psi_j \\ &= \frac{1}{2}(1-\gamma_5)\mathrm{e}^{-\mathrm{i}\alpha_j}\psi_j \\ &= \mathrm{e}^{-\mathrm{i}\alpha_j}\psi_{\mathrm{L}j} \end{aligned} \tag{9.12.35}$$

Similarly the chiral transformation of the right-handed field can be also obtained

$$\psi_{\mathrm{R}j} \to \mathrm{e}^{\mathrm{i}\alpha_j}\psi_{\mathrm{R}j} \tag{9.12.36}$$

Using Eqs. (9.12.35) and (9.12.36), we can derive the chiral transformation of the quark mass term

$$\begin{aligned}& m_j\bar{\psi}_{\mathrm{L}j}\psi_{\mathrm{R}j} + m_j^*\bar{\psi}_{\mathrm{R}j}\psi_{\mathrm{L}j} \\ &\quad \to m_j\mathrm{e}^{2\mathrm{i}\alpha_s}\bar{\psi}_{\mathrm{L}j}\psi_{\mathrm{R}j} + m_j^*\mathrm{e}^{-2\mathrm{i}\alpha_s}\bar{\psi}_{\mathrm{R}j}\psi_{\mathrm{L}j}\end{aligned} \tag{9.12.37}$$

which indicates that the mass term violates chiral transformation.

Equation (9.12.37) also shows that, as the quark field transforms under chiral transformation, if the quark mass also changes by a complex phase

$$m_j \to m_j\mathrm{e}^{-2\mathrm{i}\alpha_j} \tag{9.12.38}$$

then under the joint transformation

$$\psi_j \to \mathrm{e}^{\mathrm{i}\alpha_j\gamma_5}\psi_j, \quad m_j \to m_j\mathrm{e}^{-2\mathrm{i}\alpha_j}, \quad j = 1,\ldots,n \tag{9.12.39}$$

the QCD Lagrangian remains invariant, i.e. the joint chiral transformation for the quark field and phase transformation for quark mass are the symmetric transformation of QCD. Next we shall see that the axial vector current anomaly can violate such a symmetry [73].

Let us consider the change of the action of QCD under the transformation in Eq. (9.12.39)

$$\delta S = \delta \int \mathrm{d}^4 x \mathcal{L}_{\mathrm{QCD}} \tag{9.12.40}$$

According to the calculation in Sec. 2.2 of Chapter 2, i.e. the result of Eq. (2.2.18), and considering the mass transformation in Eq. (9.12.39), we can obtain

$$\begin{aligned} \delta S = \int \mathrm{d}^4 x \sum_j \Bigg[\partial_\mu \left(\frac{\partial \mathcal{L}_{\mathrm{QCD}}}{\partial(\partial_\mu \psi_{\mathrm{j}})} \right) \frac{\delta \psi_j}{\delta \alpha_j} + \frac{\partial \mathcal{L}_{\mathrm{QCD}}}{\partial m_{\mathrm{j}}} \frac{\delta m_j}{\delta \alpha_j} \\ + \frac{\partial \mathcal{L}_{\mathrm{QCD}}}{\partial m_j^*} \frac{\delta m_j^*}{\delta \alpha_j} \Bigg] \Bigg|_{\alpha_j \to 0} \alpha_j = \int \mathrm{d}^4 x \sum_j [\partial_\mu (\bar{\psi}_j \mathrm{i} \gamma_\mu \mathrm{i} \gamma_5 \psi_j) \\ - (-2\mathrm{i} m_j) \bar{\psi}_{Lj} \psi_{Rj} - (-2\mathrm{i} m_j)^* \bar{\psi}_{Rj} \psi_{Lj}] \alpha_j \end{aligned} \tag{9.12.41}$$

As the m_j after transformation is real, we have

$$\delta S = \int \mathrm{d}^4 x \sum_j [-\partial_\mu (\bar{\psi}_j \gamma_\mu \gamma_5 \psi_j) + 2\mathrm{i} m_j \bar{\psi}_j \gamma_5 \psi_j] \alpha_j \tag{9.12.42}$$

According to the axial current anomaly [74, 75],

$$\partial_\mu (\bar{\psi}_j \gamma_\mu \gamma_5 \psi_j) = 2\mathrm{i} m_j \bar{\psi}_j \gamma_5 \psi_j + \frac{\alpha_s}{4\pi} F^a_{\mu\nu} \tilde{F}^{a\mu\nu} \tag{9.12.43}$$

where the second term is the anomaly term, we have

$$\begin{aligned} \delta S &= - \int \mathrm{d}^4 x \sum_j \left[\frac{\alpha_s}{4\pi} F^a_{\mu\nu} \tilde{F}^{a\mu\nu} \right] \alpha_j \\ &= - \sum_j 2\alpha_j \int \mathrm{d}^4 x \frac{\alpha_s}{8\pi} F^a_{\mu\nu} \tilde{F}^{a\mu\nu} \end{aligned} \tag{9.12.44}$$

The above result indicates that the chiral symmetry of QCD is violated by the anomaly, which is true even if the quark mass is zero [73].

Equation (9.12.44) is equivalent to adding a term in the Lagrangian of QCD

$$- \sum_j 2\alpha_j \frac{\alpha_s}{8\pi} F^a_{\mu\nu} \tilde{F}^{a\mu\nu} \tag{9.12.45}$$

So under the joint chiral transformation and the quark mass phase transform, the effective Lagrangian of QCD becomes

$$\mathcal{L}_{\text{QCD}} = -\frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} + \sum_{j=1}^{n}\left(\bar{\psi}_j \mathrm{i}\gamma^\mu D_\mu \psi_j - m_j \bar{\psi}_{Lj}\psi_{Rj} - m_j^* \bar{\psi}_{Rj}\psi_{Lj}\right) + \bar{\theta}\frac{\alpha_s}{8\pi}F^a_{\mu\nu}\tilde{F}^{a\mu\nu} \tag{9.12.46}$$

where the new parameter

$$\bar{\theta} = \theta - \sum_j 2\alpha_j \tag{9.12.47}$$

The strong CP violating term may contribute to the electric dipole moment (EDM) of neutron. Previous calculation [76, 77] shows that the electric dipole moment of neutron

$$d_{\text{n}} \sim 5 \times 10^{-16}\bar{\theta} e\,\text{cm} \tag{9.12.48}$$

Current experimental measurements of neutron electric dipole moment indicate that [78]

$$d_{\text{n}} < 2.9 \times 10^{-26} e\,\text{cm} \tag{9.12.49}$$

which gives the experimental restriction on $\bar{\theta}$

$$\bar{\theta} < 0.6 \times 10^{-10} \tag{9.12.50}$$

The problem of CP violation in strong interaction is the problem why the parameter of the CP violating term for strong interaction is so small [79–81]. The transformation in Eq. (9.12.38) indicates that the quark mass must have non-zero complex phase before chiral transformation to guarantee the quark mass be real after chiral transformation. In the electromagnetic unification model of Glashow, Weinberg and Salam, quark masses are generated by Yukawa coupling between quarks and higgs field. The Yukawa coupling constants can be complex numbers in principle. The phase angles of these complex coupling constants finally lead to the complex phase angles of quark masses $-2\alpha_j$. The chiral transformation in Eq. (9.12.39) is necessary to get the quark masses to be real theoretically. Suppose that the original quark masses generated by Yukawa coupling are m'_j, and there are complex phase angles $-2\alpha_j$ for m'_j before chiral transformation, then a transformation

$$m'_j = m_j \mathrm{e}^{-2\mathrm{i}\alpha_j} \tag{9.12.51}$$

is necessary to obtain real quark masses. So the second term in Eq. (9.12.47) $-\sum_j 2\alpha_j$ is actually the complex phase angle of the determinant of the original quark mass matrix

$$-\sum_j 2\alpha_j = \arg(\det M') \tag{9.12.52}$$

where M' is the quark mass matrix before chiral transformation. Therefore, Eq. (9.12.47) can also be written as

$$\bar{\theta} = \theta + \arg(\det M') \tag{9.12.53}$$

The question now is why the phase angle of the quark mass matrix almost exactly cancels the parameter θ, thereby resulting in the parameter $\bar{\theta}$ being very small. Remember that θ comes from strong interaction, while the mass matrix comes from Yukawa coupling and weak symmetry breaking [82–84], which are essentially two different interactions.

In 1977, R.D. Peccei and H.R. Quinn provided an explanation [79, 80]. Based on the instanton effect [85–88], they proved that the CP violating parameter $\bar{\theta}$ in the Lagrangian of QCD tends to zero naturally, so that QCD remains to be CP invariant, as long as the Lagrangian of QCD keeps chiral U(1) symmetry, and at least the mass of one type of quarks generated by Yukawa coupling with a scalar field, which has a non-zero vacuum expectation value.

For the sake of brevity, let us take the quark field with only one flavor as an example and leave out the details for the proof. Suppose that a scalar field ϕ interacts with quark field in the form of Yukawa coupling, which respects the chiral U(1) symmetry. Then the Lagrangian of QCD becomes

$$\begin{aligned}\mathcal{L}_{\mathrm{QCD}} = &-\frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} + \mathrm{i}\bar{\psi}D_\mu\gamma^\mu\psi + \theta\frac{\alpha_s}{8\pi}F^a_{\mu\nu}\tilde{F}^{a\mu\nu} \\ &+ \bar{\psi}\left[G\phi\left(\frac{1+\gamma_5}{2}\right) + G^*\phi^*\left(\frac{1-\gamma_5}{2}\right)\right]\psi \\ &- |\partial_\mu\phi|^2 + \mu^2|\phi|^2 - h|\phi|^4\end{aligned} \tag{9.12.54}$$

where G is the Yukawa coupling constant, μ and h are the parameters of the scalar self-interaction potential, and $\mu^2 > 0$. $\mathcal{L}_{\mathrm{QCD}}$ keeps invariant under the following chiral transformation:

$$\psi \to \mathrm{e}^{\mathrm{i}\alpha\gamma_5}\psi, \quad \phi \to \mathrm{e}^{-2\mathrm{i}\alpha}\phi \tag{9.12.55}$$

This symmetry of QCD Lagrangian is called Peccei–Quinn symmetry. The transformation in Eq. (9.12.55) is referred to as Peccei–Quinn transformation, which is denoted as $\mathrm{U}(1)_{\mathrm{PQ}}$.

It can be seen from the last two terms of Eq. (9.12.54) that the scalar field ϕ has non-zero vacuum expectation value. Assume that the vacuum expectation value of ϕ is $\upsilon\mathrm{e}^{\mathrm{i}\beta}$, where β is the possible complex phase angle for the vacuum expectation value. Change the scalar field to be

$$\phi = \mathrm{e}^{\mathrm{i}\beta}(\upsilon + \rho + \mathrm{i}\sigma) \tag{9.12.56}$$

where ρ and σ are new fields after the change. The vacuum expectation values of ρ and σ are all zero, i.e.

$$\langle 0\,|\hat{\rho}|\;0\rangle = 0, \quad \langle 0\,|\hat{\sigma}|\;0\rangle = 0 \tag{9.12.57}$$

Here $\hat{\rho}$ and $\hat{\sigma}$ are the relevant operators of the scalar fields ρ and σ. Using the condition that the vacuum expectation values of ρ and σ are zero, Peccei and Quinn proved that there naturally is

$$\theta - 2\alpha = 0 \tag{9.12.58}$$

after quark obtains mass. The summation in Eq. (9.12.47) is not necessary at present since only one type of quark flavor is considered here, i.e.

$$\bar{\theta} = \theta - 2\alpha \tag{9.12.59}$$

This implies

$$\bar{\theta} = 0 \tag{9.12.60}$$

The generalization to multiple quark flavors and numerous scalars is straightforward. The proof of this conclusion is omitted here, and the interested readers may refer to the original literature [79, 80].

Weinberg and Wilczek proposed a theory of axion on the basis of Peccei–Quinn mechanism, where the axion is a pseudoscalar particle [89, 90]. The original axion theory rooted in the two-Higgs-doublet model [91], where the theory remains $\mathrm{U}(1)_{\mathrm{PQ}}$ symmetry. Both of the two Higgs doublets transform according to $\mathrm{U}(1)_{\mathrm{PQ}}$ transformation. One of the doublet generates mass for up type quarks by Yukawa coupling, and the other generates masses for down type quarks and leptons. The coupling between axion and fermion in this theory eventually is in the following form

$$\mathcal{L}_{\mathrm{A}\mathrm{f}\bar{\mathrm{f}}} = \frac{m_{\mathrm{f}}}{f_{\mathrm{A}}}\bar{\psi}_{\mathrm{f}}\gamma_5\psi_{\mathrm{f}}\phi_{\mathrm{A}} \tag{9.12.61}$$

where ψ_{f} and ϕ_{A} are fields of fermion and axion, respectively, m_{f} the fermion mass, and f_{A} is the parameter reflecting the strength of the coupling. In the theory of axion with two Higgs doublets, f_{A} is at the order of the electromagnetic symmetry breaking scale

$$f_{\mathrm{A}} \sim v_{\mathrm{weak}} = (\sqrt{2}G_{\mathrm{F}})^{1/2} = 246.3\,\mathrm{GeV} \tag{9.12.62}$$

and the mass of axion is estimated to be

$$m_{\mathrm{A}} \approx 100\,\mathrm{keV} \times (10^{-1}\text{–}10^{+1}) \tag{9.12.63}$$

Therefore, the axion is a particle with a very small mass.

Later more theoretical models of axion based on $\mathrm{U}(1)_{\mathrm{PQ}}$ symmetry were proposed [92–96]. Different theoretical models predict different properties of axion. In the later theoretical models, axion mass can be as low as a few eV or meV (10^{-3} eV). The exploration for axion is also carried out in various experiments, such as particle collision experiment, particle rare decay experiment, astrophysics experiment and so on [97]. But so far, none of the experiments have found evidence of axion.

References

1. Bloom E D, *et al. Phys. Rev. Lett.*, 1969, 23: 930; Breidenbach M, *et al. Phys. Rev. Lett.*, 1969, 23: 935.
2. Bjorken J D. *Phys. Rev.*, 1969, 179: 1547.
3. Feynman R P. *Phys. Rev. Lett.*, 1969, 23: 1415.
4. Gross D J, Wilczek F. *Phys. Rev. Lett.*, 1973, 30: 1343.
5. Politzer H D. *Phys. Rev. Lett.*, 1973, 30: 1346.
6. Gross D J, Wilczek F. *Phys. Rev.*, 1973, D8: 3633.
7. Gross D J, Wilczek F. *Phys. Rev.*, 1974, D9: 980.
8. Georgi H, Politzer H D. *Phys. Rev.*, 1974, D9: 416.
9. Watanabe Y, *et al. Phys. Rev. Lett.*, 1975, 35: 898.
10. Chang C, *et al. Phys. Rev. Lett.*, 1975, 35: 901.
11. Feynman R P. *Rev. Mod. Phys.*, 1948, 20: 367.
12. Schwinger J. *Proc. Natl. Acad. Sci.*, 1951, 37: 452, 455.
13. Lehmann H, Symanzik K, Zimmermann W. *Nuovo Cim.*, 1955, 1: 205.
14. Faddeev L D, Popov V N. *Phys. Lett.*, 1967, B25: 29.
15. Pauli W, Villars F. *Rev. Mod. Phys.*, 1949, 21: 434.
16. ’t Hooft G. *Nucl. Phys.*, 1971, B33: 173.

17. Ashmore J F. *Lett. Nuovo Cim.*, 1972, 4: 289.
18. Bollini C G, Giambiagi J J. *Phys. Lett.*, 1972, B40: 566.
19. Cicuta G M, Moutaldi E. *Lett. Nuovo Cim.*, 1972, 4: 329.
20. 't Hooft G, Veltman M. *Nucl. Phys.*, 1972, B44: 189.
21. Adler S L. *Phys. Rev.*, 1969, 177: 2426.
22. Bell J S, Jackiw R. *Nuovo Cim.*, 1969, A60: 47.
23. Weinberg S. *Phys. Rev.*, 1960, 118: 838.
24. Nakanishi N. *Prog. Theor. Phys.*, 1957, 17: 401; *J. Math. Phys.*, 1963, 4: 1385.
25. Hahn Y, Zimmermann W. *Comm. Math. Phys.*, 1968, 10: 330.
26. Zimmermann W. *Comm. Math. Phys.*, 1968, 11: 1.
27. 't Hooft G. *Nucl. Phys.*, 1973, B61: 455.
28. Bardeen W A, *et al. Phys. Rev.*, 1978, D18: 3998.
29. Muta T. *Foundations of Quantum Chromodynamics.* World Scientific Publishing Co. Pte. Ltd, 1998.
30. Becchi C, Rouet A, Stora R. *Ann. Phys.*, 1976, 98: 287.
31. Tyutin I V. Lebedev Institute preprint (unpublished), 1975.
32. Iofa M Z, Tyutin I V. *Theor. Math. Phys.*, 1976, 27: 316.
33. Slavnov A A. *Theor. Mat. Fiz.*, 1972, 10: 153; *Theor. Math. Phys.*, 1973, 10: 99.
34. Taylor J C. *Nucl. Phys.*, 1971, B33: 436.
35. Kluberg-Stern H, Zuber J B. *Phys. Rev.*, 1975, D12: 467.
36. Wilson K G. *Phys. Rev.*, 1971, D3: 1818.
37. Stueckelberg E C G, Petermann A. *Helv. Phys. Acta*, 1953, 26: 499.
38. Gell-Mann M, Low F E. *Phys. Rev.*, 1954, 95: 1300.
39. Ovsyannikov L V. Dokl. *Acad. Nauk SSSR*, 1956, 109: 1112.
40. Symanzik K. *Comm. Math. Phys.*, 1970, 18: 227.
41. Callan Jr. C G. *Phys. Rev.*, 1970, D2: 1541.
42. Weinberg S. *Phys. Rev.*, 1973, D8: 3497.
43. Caswell W E, Wilczek F. *Phys. Lett.*, 1974, B49: 291.
44. Gross D J. *Methods in Field Theory*, eds. Baliass R, Zinn-Justin J. North-Holland, 1976: 141.
45. Holwerda M J, van Neerven W L, van Royen R P. *Nucl. Phys.*, 1974, B75: 302.
46. Collins J C, Macfarlane A J. *Phys. Rev.*, 1974, D10: 1201.
47. Georgi H, Politzer H D. *Phys. Rev.*, 1976, D14: 1829.
48. Landau L D, Abrikosov A A, Khalatnikov I M. *Dokl. Akad. Nauk USSR*, 1954, 95: 497, 773, 1177.
49. Landau L D, Abrikosov A A, Khalatnikov I M. *Nuovo Cim. Suppl. Ser. X*, 1956, 3: 80.
50. Jones D R T. *Nucl. Phys.*, 1974, B75: 531.
51. Caswell W E. *Phys. Rev. Lett.*, 1974, 33: 244.
52. Tarasov O V, Vladimirov A A, Zharkov A Y. *Phys. Lett.*, 1980, B93: 429.
53. van Ritbergen T, Vermaseren J A M, Larin S A. *Phys. Lett.*, 1997, B400: 379.

54. Blümlein J, Kochelev N. *Nucl. Phys.*, 1997, B498: 285; Forte S, *et al. Nucl. Phys.*, 2001, B602: 585; Anselmino M, *et al. Z. Phys.*, 1994, C64: 267.
55. PDG Collab. *Phys. Rev.*, 2012, D86: 237, 010001.
56. Friedman J T, Kendall H W. *Ann. Rev. Nucl. Sci.*, 1972, 22: 203.
57. Tu T S(Dong-Sheng Du). *Phys. Lett.* 1980, B89: 237.
58. Martin A D, *et al.* (MSTW Collab.) *Euro. Phys. J.*, 2009, C63: 189.
59. PDG Collab. *Phys. Rev.*, 2012, D86: 317, 010001.
60. Eichten E, *et al. Phys. Rev.*, 1978, D17: 3090.
61. PDG Collab. *Phys. Rev.*, 2012, D86: 1146, 1223.
62. The method for determining the parity of glueballs and hybrid states is provided by Prof. Shilin Zhu.
63. Morningstar C, Peardon M. *Phys. Rev.*, 1999, D60: 034509.
64. Chen Y, *et al. Phys. Rev.*, 2006, D73: 014516.
65. Lee W J, Weingarten D. *Phys. Rev.*, 2000, D61: 014015.
66. Bali G S, *et al. Phys. Lett.*, 1993, B309: 378; Michael C. *AIP Conf. Proc.*, 1998, 432: 657.
67. Isgur N, Paton J. *Phys. Rev.*, 1985, D31: 2910.
68. Lacock P, *et al. Phys. Lett.*, 1997, B401: 308; Bernard C, *et al. Phys. Rev.*, 1997, D56: 7039; 2003, D68: 074505.
69. Chanowitz M, Sharpe S. *Nucl. Phys.*, 1983, B222: 211; Barnes T, *et al. Nucl. Phys.*, 1983, B224: 241.
70. Amsler C, Close F E. *Phys. Rev.*, 1996, D53: 295; Close F E, Kirk A. *Eur. Phys. J.*, 2001, C21: 531.
71. Amsler C, Tömqvist N A. *Phys. Rep.*, 2004, 389: 61.
72. PDG Collab. *Phys. Rev.*, 2012, D86: 225, 010001.
73. 't Hooft G. *Phys. Rev. Lett.*, 1976, 37:8.
74. Bell J S, Jackiw R. *Nuovo Cim.*, 1969, A60: 47.
75. Adler S L. *Phys. Rev.*, 1969, 117: 2426.
76. Baluni V. *Phys. Rev. D*, 1979, 19: 2227.
77. Crewther R J, *et al. Phys. Lett. B*, 1979, 88: 123.
78. Baker C A, *et al. Phys. Rev. Lett.*, 2006, 97: 131801.
79. Peccei R D, Quinn H R. *Phys. Rev. Lett.*, 1977, 38: 1440.
80. Peccei R D, Quinn H R. *Phys. Rev. D*, 1977, 16: 1791.
81. Sikivie P. *Physique*, 2012, 13: 176.
82. Cheng H Y. *Phys. Rep.*, 1988, 158: 1.
83. Gérard J M, Mertens P. *Phys. Lett. B*, 2012, 716: 316.
84. Xing Z Z. *Phys. Rev. D*, 2012, 86: 113006.
85. Belavin A A, *et al. Phys. Lett. B*, 1975, 59: 85
86. 't Hooft G. *Phys. Rev.* D, 1976, 14: 3432.
87. Jackiw R, Rebbi C. *Phys. Rev. Lett.*, 1976, 37: 172.
88. Callan C G, Dashen R F, Gross D J. *Phys. Lett. B*, 1976, 63: 334.
89. Weinberg S. *Phys. Rev. Lett.*, 1978, 40: 223.
90. Wilczek F. *Phys. Rev. Lett.*, 1978, 40: 279.
91. Gunion J F, *et al. The Higgs Hunter's Guide.* Addison-Wesley Publishing Company, 1990.
92. Kim J E. *Phys. Rev. Lett.*, 1979, 43: 103.

93. Shifman M A, Vainstein A I, Zakharov V I. *Nucl. Phys. B*, 1980, 166: 493.
94. Zhitnitsky A R. *Sov. J. Nucl. Phys.*, 1980, 31: 260.
95. Dine M, Fischler W, Srednicki M. *Phys. Lett. B*, 1981, 104: 199.
96. Peccei R D, Wu T T, Yanagida T. *Phys. Lett. B*, 1986, 172: 435.
97. Patrignani C, *et al.* (Particle Data Group). *Chin. Phys. C.*, 2016, 40: 10001. review by A. Ringwald, L.J. Rosenberg and G. Rybka, Axions and other similar particles.

Chapter 10

Neutrino Oscillation

In order to explain nuclear β decay Pauli proposed the neutrino assumption in 1930 [1]. Since then neutrinos have tightly bounded with many physicists. Later, the neutrino's mystery veil was explored layer by layer. Now it is verified in experiments that [2] there exist three types of different neutrinos ν_{e}, ν_μ, ν_τ, their masses are almost zero, and their corresponding lepton numbers L_{e}, L_μ, L_τ conserve separately. Scientists have found that the interactions of neutrinos with matter are so weak that they (neutrinos) can easily penetrate the Sun, Earth, and all the planets in the Universe. Again, neutrinos play an important role in the formation and development processes of the Universe. According to the Big Bang theory, after one second from the Big Bang neutrinos decouple from matter and radiation in the Universe [3]. After long-time cooling of the Universe, photons and neutrinos will form the cosmic radiation backgrounds. Today, we already measured accurately the cosmic microwave background of photons, that is, the 3K CMB ($T_\gamma = 2.725 \pm 0.001\,\mathrm{K}$). But for the neutrino background CνB we still not measured. According to the prediction of the Hot Big Bang model [3], the temperature of CνB $T_\nu = 1.945 \pm 0.001\,\mathrm{K}$, and the corresponding neutrino energy $E_\nu = 1.6 \times 10^{-4}\,\mathrm{eV}$. The interaction cross section of so low-energy neutrino with matter is the order of $(G_{\mathrm{F}} E_\nu)^2 \sim 1.4 \times 10^{-63}\,\mathrm{cm}^2$. So small cross section makes the measurement of CνB very difficult. Although it is a big challenge, but it is very important issue for neutrino astrophysicists to investigate today. Recent experiments show that neutrinos have non-zero masses, and the three different types of neutrinos have different masses. Although their masses are very small but not zero, this will lead to neutrino oscillation, and $L_{\mathrm{e}}, L_\mu, L_\tau$ are no longer conserved rigorously.

Similar to quarks, in the standard model the interaction Lagrangian for the leptonic charged currents can be written as

$$L_{\mathrm{CC}} = \frac{g}{\sqrt{2}}\,(\bar{\mathrm{e}}, \bar{\mu}, \bar{\tau})_L\,\gamma^\mu \begin{pmatrix} \nu_{\mathrm{e}} \\ \nu_\mu \\ \nu_\tau \end{pmatrix}_L W_\mu^- + \mathrm{h.c.} \tag{10.1}$$

In Eq. (10.1), we have chosen a special basis to make the flavor eigenstates of the charged particles e, μ, τ coincides with their mass eigenstates. If neutrinos are massless, then their flavor eigenstates coincide with their mass eigenstates. But if neutrinos are massive, then their flavor eigenstates in Eq. (10.1) ν_{e}, ν_μ, ν_τ no longer coincide with their mass eigenstates ν_1, ν_2, ν_3. In that case we can write,

$$L_{\mathrm{CC}} = \frac{g}{\sqrt{2}}(\bar{\mathrm{e}}, \bar{\mu}, \bar{\tau})_L\gamma^\mu U \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix}_L W_\mu^- + \mathrm{h.c.} \tag{10.2}$$

where

$$\begin{pmatrix} \nu_{\mathrm{e}} \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = U \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix}, \quad U = \begin{pmatrix} U_{\mathrm{e}1} & U_{\mathrm{e}2} & U_{\mathrm{e}3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} \tag{10.3}$$

Equations (10.2) and (10.3) are written using particle field operators, if we use bra-ket state vectors $|\nu_\alpha\rangle$ $(\alpha = \mathrm{e}, \mu, \tau)$ and $|\nu_i\rangle$ $(i = 1, 2, 3)$, because $\nu_\alpha^+(x)$ acting on the vacuum state $|0\rangle$ creates $|\nu_\alpha\rangle$, then we should have

$$|\nu_\alpha\rangle = \nu_\alpha^+(x)|0\rangle \tag{10.4}$$

and

$$\begin{aligned} |\nu_\alpha\rangle &= \sum_i U_{\alpha i}^*(x)|\nu_i\rangle \\ |\nu_i &= \sum_\alpha U_{\alpha i}(x)|\nu_\alpha\rangle \end{aligned} \tag{10.5}$$

Note that the U matrix in Eq. (10.3) is unitary, it satisfies $U^+U = UU^+ = 1$ so as to guarantee lepton flavor conservation, i.e. ν_α interacts with matter through charged current in Eq. (10.1) and Eq. (10.2) will produce l_α $(\alpha = \mathrm{e}, \mu, \tau)$. The matrix U is called Maki–Nakagawa–Sakata lepton mixing matrix [4]. Later on we shall give out their parametrization form.

10.1. Experimental Evidence for Neutrino Oscillation

In this section, we shall list four kinds of different experimental evidences which show neutrino oscillations.

10.1.1. *Atmospheric neutrino oscillation*

The earliest measurement of atmospheric neutrino oscillation was proceeded at Japanese deep ground detector Super-Kamiokande (SK) in 1998 [5]. This is a Cherenkov counter-shaped cylindrical container filled up with 50,000 tons of water, and it is covered with more than 10,000 photoamplifiers all around the container wall. When the cosmic ray arrives at atmosphere, it will produce neutrinos and antineutrinos $\nu_{\mathrm{e}}\,(\bar{\nu}_{\mathrm{e}})$, $\nu_\mu\,(\bar{\nu}_\mu)$ through collisions with matter. These neutrinos and antineutrinos will shoot into SK detector from all directions. If there is no neutrino oscillation, all events for every flavored neutrino coming in and out of SK detector should be spherically symmetric. Certainly, the observed events of down and up SK should all be the same. But the observed results show that for $\nu_{\mathrm{e}}\,(\bar{\nu}_{\mathrm{e}})$ the down and up events are almost the same; however for $\nu_\mu\,(\bar{\nu}_\mu)$, the numbers of events for down and up directions are obviously different. Put the direction perpendicular to SK detector as the z-axis, the polar angle is Θ. In SK detector, the reaction process of a neutrino event is

$$\nu_\alpha(\bar{\nu}_\alpha) + \mathrm{N} \to l_\alpha(\bar{l}_\alpha) + \mathrm{X}, \quad \alpha = \mathrm{e}, \mu$$

$l_\alpha\,(\bar{l}_\alpha)$ are the leptons (antileptons) with the same flavors of $\nu_\alpha\,(\bar{\nu}_\alpha)$. We call $\nu_{\mathrm{e}}\,(\bar{\nu}_{\mathrm{e}})$ events as e-like events, $\nu_\mu\,(\bar{\nu}_\mu)$ events as μ-like events. $U(-1 < \cos\Theta < -0.2)$ represents the up-going event number, $D\,(0.2 < \cos\Theta < 1)$ represents the own-going event number. The result of the SK measurements is [6] as follows.

For multi-GeV e-like events, the ratio of upgoing and downgoing events is given by

$$\frac{U_{\mathrm{e}}}{D_{\mathrm{e}}} = 0.961^{+0.086}_{-0.079} \pm 0.016 \tag{10.1.1}$$

For single-ring multi-GeV μ-like events, the ratio of upgoing and downgoing events is given by

$$\frac{U_\mu}{D_\mu} = 0.551^{+0.035}_{-0.033} \pm 0.004 \tag{10.1.2}$$

In the above two equations, the first error is statistical, the second one is systematic. Here single-ring means single Cherenkov light ring event in the SK detector.

How do we explain the result of Eq. (10.1.2) theoretically? Assume that there are only two types of neutrinos ν_μ ν_τ, then Eqs. (10.3) and (10.5) become very simple:

$$\begin{aligned} \nu_\mu &= \nu_1 \cos\theta + \nu_2 \sin\theta \\ \nu_\tau &= -\nu_1 \sin\theta + \nu_2 \cos\theta \end{aligned} \tag{10.1.3}$$

In vacuum, the probability for $\nu_\alpha\,(\alpha = \mu, \tau)$ flying a distance L then changing into ν_β is

$$P(\nu_\alpha \to \nu_\beta) = \sin^2 2\theta \sin^2\left(1.27\frac{\Delta m^2(\mathrm{eV}^2)}{E(\mathrm{GeV})}L(\mathrm{km})\right) \tag{10.1.4}$$

The derivation of Eq. (10.1.4) will be given in Sec. 10.2.

We need to point out that Eq. (10.1.4) is applicable for $\nu_\mu \to \nu_\tau$ oscillation both in vacuum and matter. Using Eq. (10.1.4), we can explain the result of Eq. (10.1.2), that is, when upgoing ν_μ penetrating the earth, a part of ν_μ oscillates into ν_τ, so upgoing ν_μ decreases.

SK also measured the curve for the probability of $\nu_\mu \to \nu_\tau$ $P(\nu_\mu \to \nu_\tau)$ as a function of $\frac{L}{E}$ which is sketched in Fig. 10.1. Figure 10.1 presents the ratio of the experimental value to the Monte Carlo simulation value

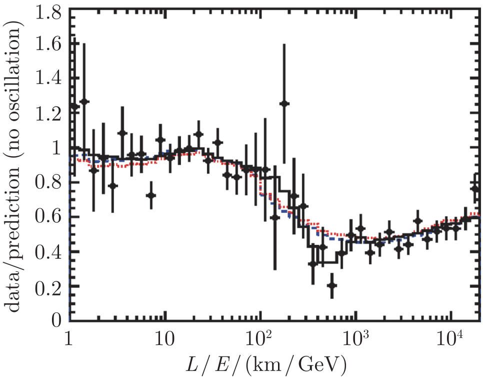


Fig. 10.1. The black points show the ratio for the experimental values to the Monte Carlo ones without oscillation as a function of L/E, the solid line is for the best-fit $\nu_\mu \to \nu_\tau$ oscillation, the dashed line is for the best-fit expectation value for neutrino decay model, the dotted line is for the best-fit expectation value for the neutrino decoherence model

without oscillation as a function of $\frac{L}{E}$ [7]. Obviously, only $\nu_\mu \to \nu_\tau$ oscillation assumption of Eq. (10.1.4) can give the dip of the experimental curve. While $\nu_\mu \to \nu_{\mathrm{e}}$ and $\nu_\mu \to \nu_{\mathrm{S}}$ (Sterile ν), oscillations can not explain the data.

From the fitted experimental points in Fig. 10.1 we can get $1.9 \times 10^{-3} < \left|\Delta m^2\right| < 3.0 \times 10^{-3}\,\mathrm{eV}^2$, $\sin^2 2\theta > 0.9$ at 90% CL.

10.1.2. *Accelerator neutrino oscillation*

As mentioned above, the number of the upgoing atmospheric neutrinos ν_μ decreases after penetrating the earth. That effect can be attributed to oscillation of ν_μ to ν_τ during the penetrating process. This is called ν_μ disappearance. This phenomenon of ν_μ disappearance should be also observed in the flying process of the accelerator produced ν_μ. Of course it needs a long enough flying distance, this is so-called long-baseline neutrino experiment. There are two experiments which observed ν_μ disappearance. one is K2K [8], its ν_μ beam produced in Japanese KEK accelerator flies 250 km away and is then detected by SK detector. The other is MINOS [9], its ν_μ beam produced in American Fermilaboratory accelerator flies 735 km away and is then detected by a calorimeter in Soudan deep mine in Minnesota. In these two experiments they all see the decreases of the number of ν_μ after flying a long distance. These decreases can all be explained by $\nu_\mu \to \nu_\tau$ oscillation.

In Fig. 10.2, the curve of the MINOS data constrain to Δm^2, $\sin^2 2\theta$ is sketched [10]. In the same figure

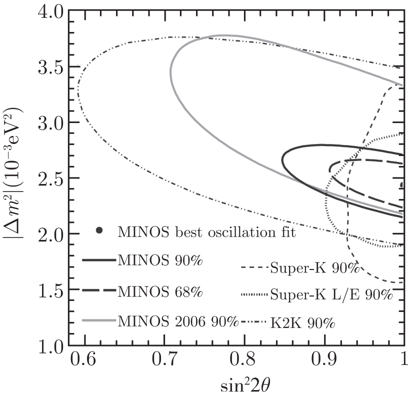


Fig. 10.2. MINOS data-fitted curve of 2008: the Super-K and K2K data fitted curves are also given.

The curves for Super-K atmospheric neutrino experiment and K2K experiment are also sketched. The best-fit of the MINOS data gives

$$\left|\Delta m^2\right| = 2.43 \times 10^{-3}\mathrm{eV}^2, \quad \sin^2 2\theta = 1.0$$

10.1.3. *Solar neutrino oscillation*

The research for the puzzle of solar neutrino disappearance has been lasting for more than 30 years. The earliest experimental research for the solar neutrino ν_e disappearance was pursued by the American Homestake Group [11]. This group put a cylindrical container filled up with 390 thousand liter liquid $\mathrm{C_2Cl_4}$ (equivalent to 520t liquid chlorine) into 4850 feet deep Homestake gold mine (South Dakota state). They used the process of $\nu_\mathrm{e} + {}^{37}\mathrm{Cl} \to {}^{37}\mathrm{Ar} + \mathrm{e}^-$ to measure ν_e. Later on other groups [12] GALLEX, GNO (Italian mountain cave Gran Sasso) and SAGE (Russia) used the process of

$$\nu_\mathrm{e} + {}^{71}\mathrm{Ga} \to {}^{71}\mathrm{Ge} + \mathrm{e}^-$$

to detect solar neutrino ν_e.

The results of these experiments show that the flux of solar neutrino ν_e received at earth is much less than that predicted by the standard solar model (SSM) [13]. The results are shown in Table 10.1 [14].

The relatively decisive experiment is done by SNO (Sudbury Neutrino Observatory) Collaboration [15]. This group used solar ^{8}B neutrino (energy E_{ν_e} is 6–7 MeV) to measure three different ν_e fluxes which are

Table 10.1. Solar neutrino flux measured using radiochemistry of all the above groups. (In Solar Neutrino Unit (SNU), SNU $= 10^{-36}\boldsymbol{\nu}_\mathrm{e}$ capture per atom per second.)

	$^{37}\mathrm{Cl} \to {}^{37}\mathrm{Ar}$	$^{71}\mathrm{Ga} \to {}^{71}\mathrm{Ge}$
Homestake	$2.56 \pm 0.16 \pm 0.16$	
GALLEX		$73.4^{+6.1}_{-6.0}\ {}^{+3.7}_{-4.1}$
GNO		$62.9^{+5.5}_{-5.3}\ \pm 2.5$
GNO+GALLEX		$67.6^{+4.0}_{-4.0}\ {}^{+3.2}_{-3.2}$
SAGE		$65.4^{+3.1}_{-3.0}\ {}^{+2.6}_{-2.8}$
SSM predicts	$8.46^{+0.87}_{-0.88}$	$127.9^{+8.1}_{-8.2}$

given through three different interactions for ν_{e} with heavy water (charged current interaction (CC), neutral current interaction (NC), and elastic scattering (ES)). Here CC interaction process is $\nu_{\mathrm{e}}\mathrm{D} \to \mathrm{e}^-\mathrm{pp}$, NC process is $\nu_\alpha \mathrm{D} \to \nu_\alpha \mathrm{pn}\,(\alpha = \mathrm{e}, \mu, \tau)$, ES process is $\nu_\alpha \mathrm{e}^- \to \nu_\alpha \mathrm{e}^-$. If there is no oscillation, when the solar neutrino ν_{e} arrives SNO detector, its flux will be no change, therefore the ν_{e} fluxes measured through the three CC, NC, ES interactions should be the same and equal to the calculated ones with the standard solar model. But the experimental results are not the case! In fact, CC events can measure only ν_{e} flux ϕ_{e}, NC events can measure all three ν_{e}, ν_μ, ν_τ fluxes, and NC interaction has nothing to do with the neutrino's flavor. We write the flux of NC events ϕ_{NC} as $\phi_{\mathrm{NC}} = \phi_{\mathrm{e}} + \phi_{\mu\tau}$, where $\phi_{\mu\tau}$ means the sum of the fluxes of ν_μ and ν_τ. ES event flux is also the total sum of the event fluxes of ν_{e}, ν_μ, ν_τ. But the cross sections of the elastic scattering of ν_μ, ν_τ with matter is less than that of ν_{e}, the suppression factor is denoted by $\sigma_\mu/\sigma_{\mathrm{e}}$, this factor is 0.155, so

$$\phi_{\mathrm{CC}} = \phi_{\mathrm{e}}, \quad \phi_{\mathrm{NC}} = \phi_{\mathrm{e}} + \phi_{\mu\tau}, \quad \phi_{\mathrm{ES}} = \phi_{\mathrm{e}} + \frac{\phi_\mu}{\phi_{\mathrm{e}}}\phi_{\mu\tau} \approx \phi_{\mathrm{e}} + 0.155\phi_{\mu\tau} \tag{10.1.5}$$

The measured result of SNO is as follows [16]:

$$\phi_{\mathrm{CC}}^{\mathrm{SNO}} = 1.68 \pm 0.06\,(\mathrm{stat.})^{+0.08}_{-0.09}\,(\mathrm{syst.}) \times 10^6\,\mathrm{cm}^{-2}\mathrm{s}^{-1} \tag{10.1.6}$$

$$\phi_{\mathrm{ES}}^{\mathrm{SNO}} = 2.35 \pm 0.02\,(\mathrm{stat.}) \pm 0.15\,(\mathrm{syst.}) \times 10^6\,\mathrm{cm}^{-2}\mathrm{s}^{-1} \tag{10.1.7}$$

$$\phi_{\mathrm{NC}}^{\mathrm{SNO}} = 4.94 \pm 0.21\,(\mathrm{stat.})^{+0.38}_{-0.34}\,(\mathrm{syst.}) \times 10^6\,\mathrm{cm}^{-2}\mathrm{s}^{-1} \tag{10.1.8}$$

$\phi_{\mathrm{NC}}^{\mathrm{SNO}}$ should be equal to the total ν_{e} flux calculated by use of the standard solar model (because it includes the fluxes of ν_{e} transformed into ν_μ and ν_τ on the way of arriving at the earth). So, the measured value of $\phi_{\mathrm{NC}}^{\mathrm{SNO}}$ in Eq. (10.1.8) is also a test for the correctness of the standard solar model. We are happy to see that the value in Eq. (10.1.8) is consistent with the predicted one by the standard solar model (SSM) [13] (SSM predicts $\phi_{\nu_{\mathrm{e}}}^{\mathrm{SSM}} = 5.05^{+1.01}_{-0.81}$ (syst.) $\times 10^6\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$).

In Fig. 10.3, the flux of $\nu_\mu + \nu_\tau$ $\phi_{\mu\tau}$ as a function of the flux of $\nu_{\mathrm{e}}\phi_{\mathrm{e}}$ is sketched. The data can be completely explained by the oscillation of $\nu_{\mathrm{e}} \to \nu_\mu$, $\nu_{\mathrm{e}} \to \nu_\tau$. From Eqs. (10.1.6), (10.1.8), and (10.1.5), we can obtain

$$\phi_{\mu\tau} = \left(3.26 \pm 0.25^{+0.40}_{-0.35}\right) \times 10^6\,\mathrm{cm}^{-2}\mathrm{s}^{-1} \tag{10.1.9}$$

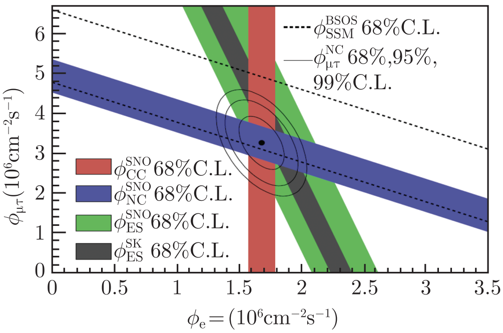


Fig. 10.3. The relation of $\phi_{\mu\tau}$ and ϕ_{e}. The broadband surrounded by the dashed line is the ν_{e} flux predicted by SSM [17], the blue broadband parallel to it is the SNO measured NC flux, the vertical red broadband is $\phi_{\mathrm{CC}}^{\mathrm{SNO}}$, the slightly inclined green broadband is $\phi_{\mathrm{ES}}^{\mathrm{SNO}}$, the black broadband in the middle of the green band is the Super-K flux $\phi_{\mathrm{ES}}^{\mathrm{SK}}$.

$\phi_{\mu\tau} \neq 0$ demonstrates that a part of ν_{e} oscillated into ν_{μ}, ν_{τ} during the process of arriving at the earth, this is the reason why a part of the solar neutrinos ν_{e} disappeared when they arrived at the earth.

10.1.4. *Reactor neutrino oscillation*

The decisive experiment was done by KamLand Group. KamLand is a neutrino detector consist of 1000t super-pure liquid scintillators which was put in Japanese Kamiokande mine. The $\bar{\nu}_{\mathrm{e}}$ beam produced by Japanese nuclear reactor flies about 180 km (weighted mean distance) and arrives at KamLand detector. The experiment measures the surviving probability $P(\bar{\nu}_{\mathrm{e}} \to \bar{\nu}_{\mathrm{e}})$ [18] for $\bar{\nu}_{\mathrm{e}}$ beam bombarding the detector through the process $\bar{\nu}_{\mathrm{e}} + \mathrm{p} \to \mathrm{e}^{+} + \mathrm{n}$.

In Fig. 10.4, the experimental curve is shown, the vertical axis is the ratio of ($\bar{\nu}_{\mathrm{e}}$ data minus background and Geo $\bar{\nu}_{\mathrm{e}}$) to $\bar{\nu}_{\mathrm{e}}$ spectrum without oscillation, the horizontal axis is L_0/E, L_0 is the $\bar{\nu}_{\mathrm{e}}$ traveled weighted mean baseline distance (180 km), while Geo $\bar{\nu}_{\mathrm{e}}$ is the $\bar{\nu}_{\mathrm{e}}$ background produced by radio chemical elements on the earth.

The curve in Fig. 10.4 shows the obvious $\sin^2\left(\frac{\Delta m^2 L}{E}\right)$ characteristic.

Note that if CPT invariance is assumed, then Eq. (10.1.4) is also applicable to antineutrino. So we can write

$$P(\bar{\nu}_{\mathrm{e}} \to \bar{\nu}_{\mu}) = \sin^2 2\theta \sin^2\left(1.27\frac{\Delta m^2}{E}L\right) \tag{10.1.10}$$

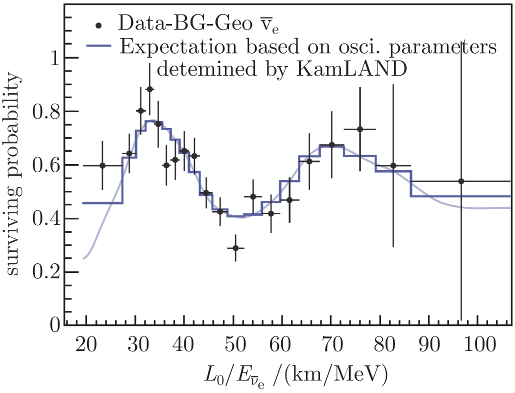


Fig. 10.4. The curve for (Data-BG-Geo-$\bar{\nu}_e$)/$\bar{\nu}_e$ spectrum(no oscillation) — $L_0/E_{\bar{\nu}_e}$

Then the $\bar{\nu}_e$ surviving probability after traveling distance L(km) is

$$P(\bar{\nu}_e \to \bar{\nu}_e) = 1 - P(\bar{\nu}_e \to \bar{\nu}_\mu) = 1 - \sin^2 2\theta \sin^2 \left(1.27\frac{\Delta m^2}{E}L\right) \tag{10.1.11}$$

The result of Fig. 10.4 can just be explained by Eq. (10.1.11), i.e. by a part of $\bar{\nu}_e$ oscillating into $\bar{\nu}_\mu$.

Because the solar neutrino oscillation is $\nu_e \to \nu_\mu, \nu_\tau$, the Δm^2 obtained from the reactor neutrino oscillation $\bar{\nu}_e \to \bar{\nu}_\mu$ should be consistent with $\Delta m^2_\odot$ from the solar neutrino oscillation. The fit to the curve given in Fig. 10.4 obtains

$$\Delta m^2_\odot = (7.59 \pm 0.21) \times 10^{-5}\,\text{eV}^2, \quad \tan^2\theta_\odot = 0.47^{+0.06}_{-0.05} \tag{10.1.12}$$

If we put the reactor neutrino oscillation data in Eq. (10.1.12) and the solar neutrino oscillation data on one Fig. 10.5, then the best fit for the two different kinds of experiments is just the large mixing angle solution of Eq. (10.1.12).

There is another reactor neutrino oscillation experiment which is done by Daya Bay Group in China [19]. This group verified not only the oscillation of $\bar{\nu}_e$ to $\bar{\nu}_\mu$, $\bar{\nu}_\tau$, but also measured the value of $\sin^2 2\theta_{13}$. This experiment will be discussed in detail in the part of neutrino CP violation.

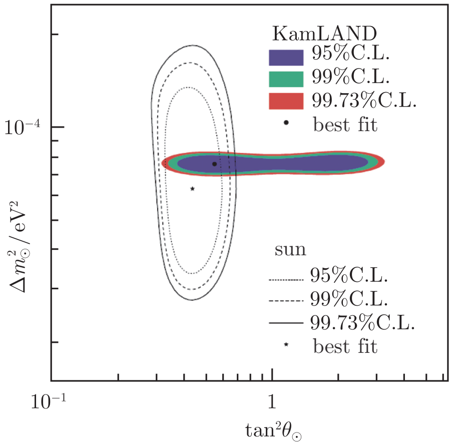


Fig. 10.5. The best-fit point for KamLand data and solar neutrino data allowed neutrino oscillation parameters $\Delta m^2_\odot$ and $\tan^2\theta_\odot$ corresponds to just the large mixing angle solution (LMA).

10.2. Theoretical Description of Neutrino Oscillation

Now we know that in nature only three different types of neutrinos are observed experimentally. Are there any other type neutrinos? We do not know. So in the following theoretical description of the neutrino oscillation we assume that there are only three different types of neutrinos.

10.2.1. *Neutrino oscillation in vacuum*

According to the definition in Eq. (10.3), the time variation of the neutrino state should be described by the mass eigenstate Schrödinger equation, i.e.

$$\mathrm{i}\frac{\partial}{\partial t}\big|\nu_i(t)\rangle = H\big|\nu_i(t)\rangle, \quad i = 1, 2, 3 \tag{10.2.1}$$

Put $|\nu_i(t=0)\rangle = |\nu_i\rangle$, then from Eq. (10.2.1) we obtain

$$|\nu_i(t)\rangle = \mathrm{e}^{-\mathrm{i}E_i t}\,|\nu_i\rangle \tag{10.2.2}$$

Because in vacuum neutrino propagates along straight line, so 1-dimensional space is enough for its description.

Assume that at $t = 0$ and $x = 0$, $|\nu_i\rangle$ state is produced, the time used for ν_i flying distance L is

$$t = L/\upsilon_i \tag{10.2.3}$$

Because the neutrino mass m_i is very small, so its velocity, $\upsilon_i = \frac{p_i}{E_i} \approx 1$, is very close to the speed of light, thus

$$t = L/\upsilon_i \approx L \tag{10.2.4}$$

When $x = L$, $t = L/\upsilon_i \approx L$, the neutrino state is given by

$$|\nu_i(t = L)\rangle = \mathrm{e}^{-\mathrm{i}E_i t}|\nu_i\rangle \approx \mathrm{e}^{-\mathrm{i}E_i L}|\nu_i\rangle \tag{10.2.5}$$

In addition

$$E_i = \sqrt{p_i^2 + m_i^2} \approx p_i \left(1 + \frac{m_i^2}{2p_i^2}\right) = p_i + \frac{m_i^2}{2p_i} \approx E + \frac{m_i^2}{2E} \tag{10.2.6}$$

Where E in Eq. (10.2.6) can be understood as the neutrino average energy, or the neutrino energy with $m_i = 0$. Here we have taken the approximation of $p_i \approx E$. In fact, from the cosmological constrain (see Sec. 10.3) the different neutrino masses are all less than 1 eV, so the approximation of Eq. (10.2.3) is very good one. For example, for the process of static π^+ meson decaying into $\mu^+\nu_\mu$, when $m_\nu \approx 0$

$$E_\nu = E = \frac{m_\pi}{2}(1 - m_\mu^2/m_\pi^2) \approx 30\,\mathrm{MeV} \tag{10.2.7}$$

When $m_\nu \neq 0$, for $\pi^+ \to \mu^+\nu_i$, according to Eq. (10.2.6)

$$E_i = E + m_i^2/2m_\pi = E(1 + 1.2 \times 10^{-16}) \tag{10.2.8}$$

$$p_i = \left|\vec{p}_i\right| \approx E(1 - 4.4 \times 10^{-16}) \tag{10.2.9}$$

E in Eq. (10.2.8) and Eq. (10.2.9) is just E in Eq. (10.2.7) with $m_\nu = 0$. Obviously this is a good approximation. Using Eq. (10.2.6), Eq. (10.2.5) can be rewritten as

$$|\nu_i(t = L)\rangle = \mathrm{e}^{-\mathrm{i}(E + \frac{m_i^2}{2E})L}\,|\nu_i\rangle \tag{10.2.10}$$

When $x = 0$, $t = 0$, according to Eq. (10.2.5) we have

$$|\nu_\alpha\rangle = \sum_i U_{\alpha i}^*\,|\nu_i\rangle \tag{10.2.11}$$

When $x = L$, $t = L$, owing to Eq. (10.2.10), $|\nu_\alpha\rangle$ becomes $|\nu_\alpha\rangle'$

$$|\nu_\alpha\rangle' = \sum_i U_{\alpha i}^*\mathrm{e}^{-\mathrm{i}(E + \frac{m_i^2}{2E})L}\,|\nu_i\rangle \tag{10.2.12}$$

Therefore, when $x = L$, $t = L$, the probability amplitude for containing $|\nu_\beta\rangle$ in $|\nu_\alpha\rangle'$ is

$$\langle\nu_\beta\ |\nu_\alpha\rangle' = \mathrm{e}^{-\mathrm{i}EL}\sum_i U_{\alpha i}^* U_{\beta i}\mathrm{e}^{-\mathrm{i}\frac{m_i^2}{2E}L} \tag{10.2.13}$$

The factor $\mathrm{e}^{-\mathrm{i}EL}$ in Eq. (10.2.13) is irrelevant with neutrino flavor, so it can be removed without affecting the probability. Thus, the probability amplitude can be simplified as

$$A(\nu_\alpha \to \nu_\beta) = \sum_i U_{\alpha i}^* U_{\beta i}\mathrm{e}^{-\mathrm{i}\frac{m_i^2}{2E}L} \tag{10.2.14}$$

The probability for ν_α after flying distance L then becoming ν_β

$$\begin{aligned} P(\nu_\alpha \to \nu_\beta) = |A(\nu_\alpha \to \nu_\beta)|^2 &= \sum_{j,k} U_{\alpha j}U_{\beta k}U_{\alpha k}^*U_{\beta j}^*\mathrm{e}^{-\mathrm{i}\frac{\Delta m_{kj}^2}{2E}L} \\ &= \sum_{j=1}^{3}\left|U_{\alpha j}U_{\beta j}^*\right|^2 + 2\,\mathrm{Re}\left\{\sum_{j<k}^{3} U_{\alpha j}U_{\beta k}U_{\alpha k}^*U_{\beta j}^*\mathrm{e}^{-\mathrm{i}\frac{\Delta m_{kj}^2}{2E}L}\right\} \end{aligned} \tag{10.2.15}$$

where

$$\Delta m_{kj}^2 = m_k^2 - m_j^2 \tag{10.2.16}$$

Eq. (10.2.15) can be also rewritten as

$$\begin{aligned} P(\nu_\alpha \to \nu_\beta) &= \sum_j \left|U_{\alpha j}U_{\beta j}^*\right|^2 + 2\sum_{j<k}\mathrm{Re}\left(U_{\alpha j}U_{\beta k}U_{\alpha k}^*U_{\beta j}^*\right)\cos\frac{\Delta m_{kj}^2}{2E}L \\ &\quad + 2\sum_{j<k}\mathrm{Im}\left(U_{\alpha j}U_{\beta k}U_{\alpha k}^*U_{\beta j}^*\right)\sin\frac{\Delta m_{kj}^2}{2E}L \\ &= \left|\sum_j U_{\alpha j}U_{\beta j}^*\right|^2 - 4\sum_{j<k}\mathrm{Re}\left(U_{\alpha j}U_{\beta k}U_{\alpha k}^*U_{\beta j}^*\right)\sin^2\frac{\Delta m_{kj}^2}{4E}L \\ &\quad + 2\sum_{j<k}\mathrm{Im}(U_{\alpha j}U_{\beta k}U_{\alpha k}^*U_{\beta j}^*)\sin\frac{\Delta m_{kj}^2}{2E}L \end{aligned}$$

$$
\begin{aligned}
&= \delta_{\alpha\beta} - 4\sum_{j<k} \mathrm{Re}(U_{\alpha j}U_{\beta k}U^*_{\alpha k}U^*_{\beta j}) \sin^2 \frac{\Delta m^2_{kj}}{4E}L \\
&\quad + 8J\sum_{\gamma} \varepsilon_{\alpha\beta\gamma} \sin\frac{\Delta m^2_{21}}{4E}L \sin\frac{\Delta m^2_{31}}{4E}L \sin\frac{\Delta m^2_{32}}{4E}L
\end{aligned} \tag{10.2.17}
$$

In Eq. (10.2.17) the unitarity of the matrix U $\sum_j U_{\alpha j}U^*_{\beta j} = \delta_{\alpha\beta}$ has been used, and the fact that the standard parameterization of the matrix U is the same as the quark mixing matrix V_{CKM} (see Eqs. (7.7.3), (7.7.5), (7.7.6)) is also used, where J is Jarlskog invariant. In the last step of Eq. (10.2.17), the following formula for the trigonometric functions is used:

$$
\begin{aligned}
\sin A \sin B \sin C &= \frac{1}{2}\left[\cos(A-B) - \cos(A+B)\right] \sin C \\
&= \frac{1}{4}[\sin(C+A-B) + \sin(C-A+B) \\
&\quad - \sin(C+A+B) - \sin(C-A-B)]
\end{aligned} \tag{10.2.18}
$$

where $A = \frac{\Delta m^2_{21}}{4E}L$, $B = \frac{\Delta m^2_{31}}{4E}L$, $C = \frac{\Delta m^2_{32}}{4E}L$.

For the oscillation probability of $\bar{\nu}_\alpha \to \bar{\nu}_\beta$, under the condition of CPT invariance $\bar{\nu}_\alpha \to \bar{\nu}_\beta$ is the CPT mirror process of $\nu_\beta \to \nu_\alpha$, i.e.

$$
P(\bar{\nu}_\alpha \to \bar{\nu}_\beta) = P(\nu_\beta \to \nu_\alpha) \tag{10.2.19}
$$

From Eq. (10.2.17)

$$
\begin{aligned}
P(\bar{\nu}_\alpha \to \bar{\nu}_\beta) &= \delta_{\alpha\beta} - 4\sum_{j<k} \mathrm{Re}(U_{\alpha j}U_{\beta k}U^*_{\alpha k}U^*_{\beta j}) \sin^2 \frac{\Delta m^2_{kj}}{4E}L \\
&\quad - 8J\sum_{\gamma} \varepsilon_{\alpha\beta\gamma} \sin\frac{\Delta m^2_{21}}{4E}L \sin\frac{\Delta m^2_{31}}{4E}L \sin\frac{\Delta m^2_{32}}{4E}L
\end{aligned} \tag{10.2.20}
$$

Equation (10.2.20) can be also obtained by making the transformations of $U \to U^*$, $J \to -J$ in Eq. (10.2.17).

When there are only two flavors of neutrinos ν_α, ν_β, there are also only two kinds of mass eigenstates ν_1, ν_2, then there is only one $\Delta m^2 = m^2_2 - m^2_1$.

The matrix U becomes

$$U = \begin{pmatrix} U_{\alpha 1} & U_{\alpha 2} \\ U_{\beta 1} & U_{\beta 2} \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \tag{10.2.21}$$

From Eq. (10.2.17), we can see that

$$\begin{aligned} P(\nu_\alpha \to \nu_\beta) &= \sin^{22}\theta \sin^2\left(\Delta m^2 \frac{L}{4E}\right) \\ &= \sin^2 2\theta \sin^2\left(1.27 \frac{\Delta m^2(\mathrm{eV}^2)}{E(\mathrm{GeV})} L(\mathrm{km})\right) \end{aligned} \tag{10.2.22}$$

This is just Eq. (10.1.4) used in Sec. 10.1.

10.2.2. *Neutrino oscillation in matter*

In this subsection, we shall use the method of analog with oscillation in vacuum to discuss oscillation in matter. So, we firstly give the standard form of oscillation in vacuum described by Hamiltonian, then extend it to the case in matter. For simplicity, we firstly discuss the case of two flavored neutrinos.

The matrix element of the Hamiltonian in vacuum H_{vac} $\langle \nu_\alpha |H_{\text{vac}}| \nu_\beta \rangle$ can be obtained from Eq. (10.5)

$$\begin{aligned} \langle \nu_\alpha |H_{\text{vac}}| \nu_\beta \rangle &= \sum_{i,j} U_{\alpha i} U^*_{\beta j} \langle \nu_i |H_{\text{vac}}| \nu_j \rangle \\ &= \sum_j U_{\alpha j} U^*_{\beta j} \sqrt{p_j^2 + m_j^2} \\ &\approx \sum_j U_{\alpha j} U^*_{\beta j} \sqrt{p^2 + m_j^2} \end{aligned} \tag{10.2.23}$$

In above, we have used $\langle \nu_i \,|\, \nu_j \rangle = \delta_{ij}$ and $H_{\text{vac}} |\, \nu_j \rangle = E_j |\, \nu_j \rangle$, and we have also used the approximation $E_j = \sqrt{p_j^2 + m_j^2} \approx \sqrt{p^2 + m_j^2}$, i.e. $p_j \approx p$, p is the momentum or energy with $m_\nu = 0$, so $E = p$.

From the discussion in Sec. 10.1 we know that when discussing flavor change (or oscillation) only the relative phase is relevant, the common phase (or the same phase) for different flavors can be dropped. Similarly the common factor for different flavors can also be dropped, because it is irrelevant to oscillation. So, in the matrix expression of H_{vac} (2×2 matrix) any factor of $E_0 \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ can be dropped.

Substituting $\sqrt{p_j^2 + m_j^2} \approx p + \frac{m_j^2}{2p}$ into Eq. (10.2.23), we get

$$\langle \nu_\alpha \left| H_{\text{vac}} \right| \nu_\beta \rangle = p\delta_{\alpha\beta} + \frac{1}{2p} \sum_j U_{\alpha j} U_{\beta j}^* m_j^2 \tag{10.2.24}$$

In Eq. (10.2.24), the first term is the form of $p\left(\begin{smallmatrix}1 & 0\\ 0 & 1\end{smallmatrix}\right)$, so it can be dropped, because it is irrelevant to oscillation. Equation (10.2.2) is simplified as

$$\langle \nu_\alpha \left| H_{\text{vac}} \right| \nu_\beta \rangle = \frac{1}{2p} \sum_j U_{\alpha j} U_{\beta j}^* m_j^2 = \frac{1}{2E} \sum_j U_{\alpha j} U_{\beta j}^* m_j^2 \tag{10.2.25}$$

Substituting the matrix U in Eq. (10.2.21) into Eq. (10.2.25), we get

$$H_{\text{vac}} = \frac{\Delta m^2}{4E} \begin{pmatrix} -\cos 2\theta + \dfrac{m_1^2 + m_2^2}{\Delta m^2} & \sin 2\theta \\ \sin 2\theta & \cos 2\theta + \dfrac{m_1^2 + m_2^2}{\Delta m^2} \end{pmatrix} \tag{10.2.26}$$

In the above equation, the constant term

$$\frac{\Delta m^2}{4E} \begin{pmatrix} \dfrac{m_1^2 + m_2^2}{\Delta m^2} & 0 \\ 0 & \dfrac{m_1^2 + m_2^2}{\Delta m^2} \end{pmatrix} = \frac{m_1^2 + m_2^2}{4E} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

can again be dropped.

Then the matrix expression of H_{vac} becomes

$$H_{\text{vac}} = \frac{\Delta m^2}{4E} \begin{pmatrix} -\cos 2\theta & \sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{pmatrix} \tag{10.2.27}$$

Equation (10.2.27) is written on the basis of $|\nu_\alpha\rangle$, $|\nu_\beta\rangle$, but because $|\nu_\alpha\rangle$, $|\nu_\beta\rangle$ are not the eigenstates of H_{vac}, $|\nu_1\rangle$, $|\nu_2\rangle$ are its eigenstates, so Eq. (10.2.27) is not diagonal. For the two flavors ν_{e} and ν_μ, according to Eq. (10.2.21) we have

$$\begin{aligned} |\nu_{\text{e}}\rangle &= |\nu_1\rangle \cos\theta + |\nu_2\rangle \sin\theta \\ |\nu_\mu\rangle &= -|\nu_1\rangle \sin\theta + |\nu_2\rangle \cos\theta \end{aligned} \tag{10.2.28}$$

Solve the eigenequation of Eq. (10.2.27), there are two eigenvalues

$$\lambda_1 = -\frac{\Delta m^2}{4E}, \quad \lambda_2 = \frac{\Delta m^2}{4E} \tag{10.2.29}$$

The time evolution of $|\nu_1\rangle, |\nu_2\rangle$ is decided by Eq. (10.2.29), so we have

$$|\nu_\mathrm{e}(t)\rangle = \mathrm{e}^{\mathrm{i}\frac{\Delta m^2}{4E}t}|\nu_1\rangle\cos\theta + \mathrm{e}^{-\mathrm{i}\frac{\Delta m^2}{4E}t}|\nu_2\rangle\sin\theta \tag{10.2.30}$$

When ν_e flies distance L, $t = L$, the probability for ν_e changing into ν_μ is

$$\begin{aligned} P(\nu_\mathrm{e}\to\nu_\mu) &= |\langle\nu_\mu|\nu_\mathrm{e}(t=L)\rangle|^2 = \left|\sin\theta\cos\theta\left(-\mathrm{e}^{\mathrm{i}\frac{\Delta m^2}{4E}L} + \mathrm{e}^{-\mathrm{i}\frac{\Delta m^2}{4E}L}\right)\right|^2 \\ &= \sin^2 2\theta\sin^2\left(\frac{\Delta m^2}{4E}L\right) \end{aligned} \tag{10.2.31}$$

This is just Eqs. (10.1.4) and (10.2.22) used in Sec. 10.1.

Now we use the method of parallel analogy to Eqs. (10.2.23)–(10.2.31) to discuss the case for ν_e penetrating matter.

According to the standard model, when ν_e penetrating matter there are charged current interactions (exchanging W) and also neutral current interactions (exchanging Z). The neutral current interactions are the same for all kinds of neutrinos of (ν_e, ν_μ, and ν_τ), but the charged current interaction of $\nu_\mathrm{e} - \mathrm{e}$ has opposite sign with that of $\bar{\nu}_\mathrm{e} - \mathrm{e}$. Thus if the effective Hamiltonian for the charged current interaction of $\nu_\mathrm{e} - \mathrm{e}$ is

$$V_\mathrm{W} = \sqrt{2}G_\mathrm{F}N_\mathrm{e} \tag{10.2.32}$$

where N_e is the electron density in unit volume of matter, then for the charged current interaction of $\bar{\nu}_\mathrm{e} - \mathrm{e}$ the corresponding effective Hamiltonian is

$$\bar{V}_\mathrm{W} = -V_\mathrm{W} = -\sqrt{2}G_\mathrm{F}N_\mathrm{e} \tag{10.2.33}$$

For the neutral current interactions, the interactions of ν_e electron e and ν_e proton p cancel each other, therefore the effective Hamiltonian for neutral current interactions of ν_e, ν_μ (through exchange of Z) is

$$V_\mathrm{Z} = -\frac{1}{\sqrt{2}}G_\mathrm{F}N_\mathrm{n} \tag{10.2.34}$$

where N_n is the neutron density of matter.

The derivation of Eqs. (10.2.32)–(10.2.34) can be found in Ref. 20. For ν_μ, its interactions with matter can only go through neutral current, so when there are only two flavor neutrinos ν_{e} and ν_μ, the effective Hamiltonian in matter H_M will be

$$H_M = H_{\mathrm{vac}} + V_{\mathrm{W}} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + V_{\mathrm{Z}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{10.2.35}$$

Because V_{Z} is the same for all flavors of neutrinos, it does not affect oscillation so can be dropped. Thus H_M can be simplified as

$$H_M = H_{\mathrm{vac}} + V_{\mathrm{W}} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = H_{\mathrm{vac}} + \frac{V_{\mathrm{W}}}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \frac{V_{\mathrm{W}}}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{10.2.36}$$

The third term in the above equation can be dropped again, we get

$$\begin{aligned} H_M &= H_{\mathrm{vac}} + \frac{V_{\mathrm{W}}}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\ &= \frac{\Delta m^2}{4E} \begin{pmatrix} -(\cos 2\theta - x) & \sin 2\theta \\ \sin 2\theta & \cos 2\theta - x \end{pmatrix} \end{aligned} \tag{10.2.37}$$

where

$$x = \frac{V_{\mathrm{W}}/2}{\Delta m^2/4E} = \frac{2\sqrt{2} G_{\mathrm{F}} N_{\mathrm{e}} E}{\Delta m^2} \tag{10.2.38}$$

Define

$$\begin{cases} \Delta m_M^2 = \Delta m^2 \sqrt{\sin^2 2\theta + (\cos 2\theta - x)^2} \\ \sin^2 2\theta_M = \dfrac{\sin^2 2\theta}{\sin^2 2\theta + (\cos 2\theta - x)^2} \end{cases} \tag{10.2.39}$$

H_M can be rewritten as

$$H_M = \frac{\Delta m_M^2}{4E} \begin{pmatrix} -\cos 2\theta_M & \sin 2\theta_M \\ \sin 2\theta_M & \cos 2\theta_M \end{pmatrix} \tag{10.2.40}$$

We can see that in Eq. (10.2.27) if we change $\theta \to \theta_M$, $\Delta m^2 \to \Delta m_M^2$, then we get Eq. (10.2.40). Obviously, the eigenstates in matter $|\nu_1\rangle_M$ and $|\nu_2\rangle_M$ are different from the eigenstates in vacuum $|\nu_1\rangle$, $|\nu_2\rangle$, and the mixing angle $\theta_M \neq \theta$. From Eq. (10.2.40), the probability of $\nu_{\mathrm{e}} \to \nu_\mu$ in matter is

$$P_M\left(\nu_{\mathrm{e}} \to \nu_\mu\right) = \sin^2 2\theta_M \sin^2\left(\frac{\Delta m_M^2}{4E}L\right) \tag{10.2.41}$$

From Eq. (10.2.39), we can see that if $\sin 2\theta \ll 1$, and $x \sim \cos 2\theta$, then $\sin^2 2\theta_M$ can be large (for instance close to 1), this case is so-called resonance.

Mikheyev–Smirnov–Wolfenstein [21] are the first to discuss the resonance effect, so this is also called MSW effect.

Above we have only discussed the simplest case and the method is taken from Ref. 22. The case for three different flavors of neutrinos can be found in [20].

Note that the parameter x in Eqs. (10.2.37) and (10.2.38) has opposite sign for ν_{e} and $\bar{\nu}_{\mathrm{e}}$. This will lead to the probability of ν_{e} oscillating to ν_μ, not equal to that of $\bar{\nu}_{\mathrm{e}}$ oscillating to $\bar{\nu}_\mu$, i.e. $P\left(\nu_{\mathrm{e}} \to \nu_\mu; t\right) \neq P(\bar{\nu}_{\mathrm{e}} \to \bar{\nu}_\mu; t)$. But this does not mean CP violation. This is the asymmetry caused by matter effects, this part of asymmetry should be deducted when discussing CP asymmetry in matter.

10.3. Dirac Neutrino and Majorana Neutrino

10.3.1. *Dirac neutrino*

In the standard $\mathrm{SU(2)_L \times U(1)_Y}$ electroweak unified model, neutrino has only left-handed component ν_{L}, and has no right-handed component ν_{R}, so it cannot construct Dirac mass term, therefore the neutrino mass is equal to zero accurately. If we extend the standard model slightly, introducing SU(2) singlet right-handed neutrino component N_{R}, and introducing Yukawa coupling

$$\begin{aligned} -L_{\mathrm{Dirac}} &= \bar{l}_{\mathrm{L}} Y_\nu \tilde{H} N_{\mathrm{R}} + \mathrm{h.c.} \\ &= \bar{l}_{\alpha\mathrm{L}} Y_\nu^{\alpha j} \tilde{H} N_{j\mathrm{R}} + \mathrm{h.c.} \end{aligned} \tag{10.3.1}$$

where $\alpha = \mathrm{e}, \mu, \tau$, $j = 1, 2, 3$, $l_{\mathrm{eL}} = \binom{\nu_{\mathrm{e}}}{\mathrm{e}}_{\mathrm{L}}$, $l_{\mu\mathrm{L}} = \binom{\nu_\mu}{\mu}_{\mathrm{L}}$, $l_{\tau\mathrm{L}} = \binom{\nu_\tau}{\tau}_{\mathrm{L}}$, $\tilde{H} = \mathrm{i}\sigma_2 H^*$, $H = \binom{H^+}{H^0}$, and $Y_\nu^{\alpha j}$ is the Yukawa coupling constant.

When $\mathrm{SU(2)_L \times U(1)_Y}$ is spontaneously broken into $\mathrm{U(1)_{em}}$, $\langle\tilde{H}\rangle = \binom{v}{0}$, then Eq. (10.3.1) becomes

$$-L'_{\mathrm{Dirac}} = \bar{\nu}_{\mathrm{L}} M_{\mathrm{D}} N_{\mathrm{R}} + \mathrm{h.c.} \tag{10.3.2}$$

where

$$\nu_{\mathrm{L}} = \begin{pmatrix} \nu_{\mathrm{eL}} \\ \nu_{\mu\mathrm{L}} \\ \nu_{\tau\mathrm{L}} \end{pmatrix}, \quad N_{\mathrm{R}} = \begin{pmatrix} N_{1\mathrm{R}} \\ N_{2\mathrm{R}} \\ N_{3\mathrm{R}} \end{pmatrix}, \quad M_{\mathrm{D}} = \upsilon Y_\nu = \upsilon \begin{pmatrix} Y_\nu^{\mathrm{e}1} & Y_\nu^{\mathrm{e}2} & Y_\nu^{\mathrm{e}3} \\ Y_\nu^{\mu 1} & Y_\nu^{\mu 2} & Y_\nu^{\mu 3} \\ Y_\nu^{\tau 1} & Y_\nu^{\tau 2} & Y_\nu^{\tau 3} \end{pmatrix} \tag{10.3.3}$$

M_{D} is the neutrino mass matrix which can be diagonalized by the two unitary matrices U, V,

$$U^+ M_{\mathrm{D}} V = M_\nu = \mathrm{diag}(m_1, m_2, m_3) \tag{10.3.4}$$

where m_i $(i = 1, 2, 3)$ are just the masses of the neutrino mass eigenstates ν_1, ν_2, ν_3. We can see in what follows, U is just the lepton mixing matrix introduced in Eq. (10.3.3).

In fact, after M_{D} is diagonalized, Eq. (10.3.2) can be rewritten as

$$\begin{aligned} -L'_{\mathrm{Dirac}} &= \bar{\nu}_{\mathrm{L}} M_{\mathrm{D}} N_{\mathrm{R}} + \mathrm{h.c.} \\ &= \bar{\nu}_{\mathrm{L}} U M_\nu V^+ N_R + \mathrm{h.c.} \\ &= \bar{\nu}'_{\mathrm{L}} M_\nu N'_{\mathrm{R}} + \mathrm{h.c.} \end{aligned} \tag{10.3.5}$$

where

$$\nu'_{\mathrm{L}} = U^+ \nu_{\mathrm{L}}, \quad N'_{\mathrm{R}} = V^+ N_{\mathrm{R}} \tag{10.3.6}$$

Put

$$\nu' = \nu'_{\mathrm{L}} + N'_{\mathrm{R}} = \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} \tag{10.3.7}$$

Then, we have

$$-L'_{\mathrm{Dirac}} = \bar{\nu}' M_\nu \nu' = \sum_{i=1}^{3} m_i \cdot \bar{\nu}_i \cdot \nu_i \tag{10.3.8}$$

ν_{L} in Eq. (10.3.1) according to Eq. (10.3.6) is just equal to

$$U \nu'_{\mathrm{L}} = U \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix}_{\mathrm{L}}$$

Note that N_{R} is SU(2) singlet, then

$$L_{\mathrm{CC}} = \frac{g}{\sqrt{2}}(\bar{\mathrm{e}}\ \bar{\mu}\ \bar{\tau})_{\mathrm{L}}\gamma^{\mu}U\begin{pmatrix}\nu_1\\ \nu_2\\ \nu_3\end{pmatrix}_{\mathrm{L}} W_{\mu}^{-} + \mathrm{h.c.} \tag{10.3.9}$$

This is just Eq. (10.3.2), so the U matrix in Eqs. (10.3.4) and (10.3.6) is just the lepton mixing matrix introduced in Eq. (10.3.3).

10.3.2. *Majorana neutrino*

Put ν_i $(i = 1, 2, 3)$ as neutrino mass eigenstates, then neutrinos which satisfy the following condition [23]

$$\nu_i^C \equiv C\bar{\nu}_i^{\mathrm{T}} = \nu_i \tag{10.3.10}$$

are called Majorana neutrinos.

From Eq. (10.3.10), we see that the antiparticle of a Majorana neutrino is just itself. The mass term of Majorana neutrinos can be written as

$$-L'_{\mathrm{Majorana}} = \frac{1}{2}\bar{\nu}_{\mathrm{L}}M(\nu_{\mathrm{L}})^C + \mathrm{h.c.} \tag{10.3.11}$$

Using the property of the charge conjugation matrix $C = \mathrm{i}\gamma^2\gamma^0$ we can prove that

$$(\nu_{\mathrm{L}})^C = (\nu^C)_{\mathrm{R}}, \quad (\nu_{\mathrm{R}})^C = (\nu^C)_{\mathrm{L}} \tag{10.3.12}$$

Because the right-hand side of Eq. (10.3.11) is a Lorentz scalar, so we have

$$\begin{aligned}\bar{\nu}_{\mathrm{L}}M(\nu_{\mathrm{L}})^C &= \left[\bar{\nu}_{\mathrm{L}}M(\nu_{\mathrm{L}})^C\right]^{\mathrm{T}} = -\bar{\nu}_{\mathrm{L}}C^T M^T\left(\bar{\nu}_{\mathrm{L}}\right)^{\mathrm{T}}\\ &= \bar{\nu}_{\mathrm{L}}M^{\mathrm{T}}(\bar{\nu}_{\mathrm{L}})^C\end{aligned} \tag{10.3.13}$$

In the above equation, the minus sign comes from the exchange of fermion fields and the equality of $C^{\mathrm{T}} = -C$. Therefore, we have $M^{\mathrm{T}} = M$, i.e. M is a symmetric matrix, so it can be diagonalized by use of a unitary matrix U

$$U^{+}MU^{*} = \hat{M}_{\nu} = \mathrm{diag}(m_1, m_2, m_3) \tag{10.3.14}$$

Equation (10.3.11) can be rewritten as

$$-L'_{\mathrm{Majorana}} = \frac{1}{2}\bar{\nu}'_{\mathrm{L}}\hat{M}_{\nu}(\nu'_{\mathrm{L}})^{C} + \mathrm{h.c.} \tag{10.3.15}$$

where

$$\nu'_{\mathrm{L}} = U^{+}\nu_{\mathrm{L}}, \quad (\nu'_{\mathrm{L}})^{C} = C\bar{\nu}'^{\mathrm{T}}_{\mathrm{L}}$$

$$\nu' = \nu'_{\mathrm{L}} + (\nu'_{\mathrm{L}})^{C} = \nu'_{\mathrm{L}} + (\nu'^{C})_{\mathrm{R}} = \nu'_{\mathrm{L}} + \nu'_{\mathrm{R}} = \begin{pmatrix}\nu_1\\ \nu_2\\ \nu_3\end{pmatrix} \tag{10.3.16}$$

At the last step in above equation, we have used Eq. (10.3.10) which gives $\nu'^{C} = \nu'$ (i.e. $\nu_i^{C} = \nu_i$).

Equation (10.3.15) can be further simplified as

$$-L'_{\mathrm{Majorana}} = \frac{1}{2}\bar{\nu}'\hat{M}_{\nu}\nu' = \frac{1}{2}\sum_{i=1}^{3} m_i \cdot \bar{\nu}_i \cdot \nu_i \tag{10.3.17}$$

The standard model which has only Higgs doublet cannot accommodate the Majorana mass term of Eq. (10.3.17), because this mass term breaks gauge invariance of $\mathrm{SU(2)_L} \times \mathrm{U(1)_Y}$. But if we introduce $\mathrm{SU(2)_L}$ Higgs triplet into the standard model, then the new model can accommodate Majorana neutrino.

Obviously, Majorana mass term breaks lepton number conservation, because it changes the lepton number by two units, i.e. $\Delta L = 2$. The factor of 1/2 in Eqs. (10.3.11) and (10.3.15) is for keeping ν_i to satisfy the Dirac equation. For the detail see Ref. [20].

The charged current interaction Lagrangian for Dirac neutrinos is expressed in Eq. (10.3.9), while for Majorana neutrinos their L_{CC} has the same form as Dirac neutrinos, but we need to replace the lepton mixing matrix in Eq. (10.3.9) by Majorana neutrino mixing matrix, and a factor of $P' = \mathrm{diag}(\mathrm{e}^{\mathrm{i}\rho}, \mathrm{e}^{\mathrm{i}\sigma}, 1)$ needs to be added. The detailed form will be given in Sec. 10.5.

The discrimination of Dirac neutrino from Majorana neutrino is very difficult task, because both participate weak interactions through the same flavor eigenstates $(\nu_{e\mathrm{L}}, \nu_{\mu\mathrm{L}}, \nu_{\tau\mathrm{L}})$. The only experiments for distinguishing the two types of neutrinos we know now are neutrinoless double β decays ($0\nu2\beta$ decays), i.e. nuclear transitions $(A, Z) \to (A, Z+2) + 2\mathrm{e}^{-}$. As sketched in Fig. 10.6, these processes change the lepton number by two units. At present, only Majorana mass term induced $\bar{\nu}_i - \nu_i$ mixing can do the job.

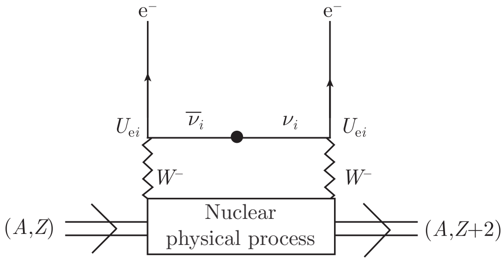


Fig. 10.6. Process for even–even nuclei exchanging virtual Majorana neutrinos induces $0\nu 2\beta$ decay.

10.4. Neutrino Mass and See-Saw Mechanism

Experiments already show that the masses of the three flavor neutrinos are all very small. The cosmological observations give that [24]

$$m_{\mathrm{e}} + m_\mu + m_\tau < 0.28\mathrm{eV}$$

The more strict constrain for the mass of $\bar{\nu}_{\mathrm{e}}$ is [25]

$$m_{\bar{\nu}_{\mathrm{e}}} < 2.05\mathrm{eV},\ \ 95\% C.L.$$

How can we guarantee so tiny neutrino masses? If the neutrino is a Dirac type, in general its Dirac mass should be the same order of magnitude as the other leptons of e, μ, τ. So it is hard to understand why the neutrino Dirac mass is so small. In what follows, we shall discuss one mechanism — the so-called See-Saw mechanism which can give small neutrino masses. The basic idea is that under the condition of gauge symmetry and Lorentz invariance we introduce, in the standard model, some new particles and allow the breakdown of B–L. At present, there are three different kinds of See-Saw mechanisms. In this section, we shall discuss one of them, i.e. Type-I See-Saw [26]. The other two can be found in [27].

Now we first discuss the case for only one generation of leptons $(\nu_{\mathrm{e}}, \mathrm{e})$. For that case, we only need to introduce one $\mathrm{SU}(2)_{\mathrm{L}}$ singlet right-handed Majorana neutrino N_{R}:

$$-L_{\mathrm{lepton}} = \bar{l}_{\mathrm{L}} Y_l H e_{\mathrm{R}} + \bar{l}_{\mathrm{L}} Y_\nu \sim H N_{\mathrm{R}} + \frac{1}{2} \bar{N}_{\mathrm{R}}^C M_{\mathrm{R}} N_{\mathrm{R}} + \mathrm{h.c.} \qquad (10.4.1)$$

where

$$\begin{cases} l_\mathrm{L} = \begin{pmatrix} \nu_\mathrm{e} \\ \mathrm{e} \end{pmatrix}_\mathrm{L}, \quad \tilde{H} = \mathrm{i}\sigma_2 H^*, \quad H = \begin{pmatrix} H^+ \\ H^0 \end{pmatrix} \\ N_\mathrm{R}^C = C\bar{N}_\mathrm{R}^\mathrm{T} \end{cases} \tag{10.4.2}$$

The last term in Eq. (10.4.1) is the Majorrana mass term of N_R, the first term is the electron mass term, the second term is the ν_e Dirac mass term. When Higgs takes its vacuum expectation value, $\langle H\rangle = \left(\begin{smallmatrix} 0 \\ v \end{smallmatrix}\right)$, $\langle \tilde{H}\rangle = \left(\begin{smallmatrix} v \\ 0 \end{smallmatrix}\right)$. We are interested in neutrino mass terms, so we take out neutrino mass part of Eq. (10.4.1) as follows:

$$-L'_\mathrm{Neutrino} = \bar{\nu}_\mathrm{eL} M_\mathrm{D} N_\mathrm{R} + \frac{1}{2}\bar{N}_\mathrm{R}^C M_\mathrm{R} N_\mathrm{R} + \mathrm{h.c.} \tag{10.4.3}$$

Note that in Eqs. (10.4.1) and (10.4.3) N_R Majorana mass term breaks lepton number conservation but keeps $\mathrm{SU}(2)_\mathrm{L} \times \mathrm{U}(1)_\mathrm{Y}$ gauge invariance.

Using

$$\bar{N}_\mathrm{R}^C M_\mathrm{D}^\mathrm{T} \nu_\mathrm{L}^C = [N_\mathrm{R}^\mathrm{T} C M_\mathrm{D}^\mathrm{T} C \bar{\nu}_\mathrm{L}^\mathrm{T}]^\mathrm{T} = \bar{\nu}_\mathrm{L} M_\mathrm{D} N_\mathrm{R} \tag{10.4.4}$$

Eq. (10.4.3) can be rewritten as

$$-L'_\mathrm{Neutrino} = \frac{1}{2}(\bar{\nu}_\mathrm{eL}\bar{N}_\mathrm{R}^C)\begin{pmatrix} 0 & M_\mathrm{D} \\ M_\mathrm{D}^\mathrm{T} & M_\mathrm{R} \end{pmatrix}\begin{pmatrix} \nu_\mathrm{eL}^C \\ N_\mathrm{R} \end{pmatrix} + \mathrm{h.c.} \tag{10.4.5}$$

Equation (10.4.5) is also applicable to the case of three lepton generations, in that case we should replace ν_eL by ν_L and

$$\nu_\mathrm{L} = \begin{pmatrix} \nu_\mathrm{eL} \\ \nu_{\mu\mathrm{L}} \\ \nu_{\tau\mathrm{L}} \end{pmatrix}, \quad N_\mathrm{R} = \begin{pmatrix} N_{1\mathrm{R}} \\ N_{2\mathrm{R}} \\ N_{3\mathrm{R}} \end{pmatrix} \tag{10.4.6}$$

where $M_\mathrm{D}, M_\mathrm{D}^\mathrm{T}, M_\mathrm{R}$ are all 3×3 matrices.

For the case of only one generation of leptons, M_D, $M_\mathrm{D}^\mathrm{T} = M_\mathrm{D}$, and M_R are all C number, the mass matrix $\left(\begin{smallmatrix} 0 & M_\mathrm{D} \\ M_\mathrm{D} & M_\mathrm{R} \end{smallmatrix}\right)$ can be diagonalized as

$$\begin{pmatrix} -2M_\mathrm{D}^2/M_\mathrm{R} & 0 \\ 0 & M_\mathrm{R} + M_\mathrm{D}^2/M_\mathrm{R} \end{pmatrix} \tag{10.4.7}$$

The above minus sign is easy to be canceled through changing the phases of ν_{eL}, ν_{eL}^C, so the diagonalized mass matrix is

$$\begin{pmatrix} 2M_{\mathrm{D}}^2/M_{\mathrm{R}} & 0 \\ 0 & M_{\mathrm{R}} + M_{\mathrm{D}}^2/M_{\mathrm{R}} \end{pmatrix} \tag{10.4.8}$$

When we assume $M_{\mathrm{R}} \gg M_{\mathrm{D}}$, Eq. (10.4.8) can lead to very small mass of the neutrino ν'

$$M'_\nu = 2M_{\mathrm{D}}^2/M_{\mathrm{R}} \tag{10.4.9}$$

where ν' is the light neutrino mass eigenstate which is the linear combination of ν_{e} and the new particle N.

This is just the basic idea of the Type-I See-Saw mechanism which can gives small neutrino mass.

For the case of three lepton generations, we can diagonalize the neutrino mass matrix according to the method of Sec. 3.2.3 in [20]. Here for simplicity in this section we did not consider the Majorana mass term of ν_{L}, i.e. we have put $M_{\mathrm{L}} = 0$. In that case, the neutrino mass part in the Lagrangian is

$$-L'_{\mathrm{Neutrino}} = \bar{\nu}_{\mathrm{L}} M_{\mathrm{D}} N_{\mathrm{R}} + \frac{1}{2}\bar{N}_{\mathrm{R}}^C M_{\mathrm{R}} N_{\mathrm{R}} + \mathrm{h.c.} \tag{10.4.10}$$

where ν_{L} and N_{R} are the matrices with three rows and one column of Eq. (10.4.6). Using Eq. (10.4.4) we can rewrite Eq. (10.4.10) as

$$-L'_{\mathrm{Neutrino}} = \frac{1}{2}(\bar{\nu}_{\mathrm{L}}, \bar{N}_{\mathrm{R}}^C) \begin{pmatrix} 0 & M_{\mathrm{D}} \\ M_{\mathrm{D}}^{\mathrm{T}} & M_{\mathrm{R}} \end{pmatrix} \begin{pmatrix} \nu_{\mathrm{L}}^C \\ N_{\mathrm{R}} \end{pmatrix} + \mathrm{h.c.} \tag{10.4.11}$$

The 6×6 mass matrix in Eq. (10.4.11) can be diagonalized by 6×6 unitary matrix, i.e.

$$\begin{pmatrix} V & R \\ S & U \end{pmatrix}^+ \begin{pmatrix} 0 & M_{\mathrm{D}} \\ M_{\mathrm{D}}^T & M_{\mathrm{R}} \end{pmatrix} \begin{pmatrix} V & R \\ S & U \end{pmatrix}^* = \begin{pmatrix} \hat{M}_\nu & 0 \\ 0 & \hat{M}_{\mathrm{N}} \end{pmatrix} \tag{10.4.12}$$

where

$$\hat{M}_\nu = \mathrm{diag}(m_1, m_2, m_3), \quad \hat{M}_{\mathrm{N}} = \mathrm{diag}(M_1, M_2, M_3)$$

and the submatrices V, R, S, U satisfy the following unitarity conditions:

$$\begin{cases} VV^+ + RR^+ = SS^+ + UU^+ = 1 \\ V^+V + S^+S = R^+R + U^+U = 1 \\ VS^+ + RU^+ = V^+R + S^+U = 0 \end{cases} \tag{10.4.13}$$

After diagonalization, Eq. (10.4.10) becomes

$$-L'_{\text{Neutrino}} = \frac{1}{2}(\bar{\nu}'_{\text{L}}, \overline{N'^C_{\text{R}}}) \begin{pmatrix} \hat{M}_\nu & 0 \\ 0 & \hat{M}_{\text{N}} \end{pmatrix} \begin{pmatrix} \nu'^C_{\text{L}} \\ N'_{\text{R}} \end{pmatrix} + \text{h.c.} \tag{10.4.14}$$

where

$$\begin{cases} \nu'_{\text{L}} = V^+\nu_{\text{L}} + S^+(N_{\text{R}})^C, \quad N'_{\text{R}} = R^{\text{T}}(\nu_{\text{L}})^C + U^{\text{T}}N_{\text{R}} \\ \nu'^C_{\text{L}} = C\bar{\nu}'^{\text{T}}_{\text{L}}, \quad N'^C_{\text{R}} = C\bar{N}'^{\text{T}}_{\text{R}} \end{cases} \tag{10.4.15}$$

Define a new field ν':

$$\nu' = \begin{pmatrix} \nu'_L \\ (N'_{\text{R}})^C \end{pmatrix} + \begin{pmatrix} (\nu'_{\text{L}})^C \\ N'_{\text{R}} \end{pmatrix} = \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \\ N_1 \\ N_2 \\ N_3 \end{pmatrix} \tag{10.4.16}$$

then

$$-L'_{\text{Neutrino}} = \frac{1}{2}\bar{\nu}' \begin{pmatrix} \hat{M}_\nu & 0 \\ 0 & \hat{M}_{\text{N}} \end{pmatrix} \nu' = \frac{1}{2}\sum_{i=1}^{3}(m_i\bar{\nu}_i\nu_i + M_i\bar{N}_iN_i) \tag{10.4.17}$$

From Eq. (10.4.16) we can obviously have $(\bar{\nu}')^C = \nu'$, i.e. ν' are six Majorana neutrinos ν_i, N_i in Eq. (10.4.17).

Now we discuss how can neutrinos get small masses.

Similar to the case of one generation leptons, in the mass term of Eq. (10.4.11) if $M_{\text{R}} \gg M_{\text{D}}$, then the problem can be solved.

In fact, $M_{\text{D}} \sim \langle H \rangle \approx 174\,\text{GeV}$ (the energy scale of the electroweak symmetry breaking). While $M_{\text{R}} \gg \langle H \rangle$ is a reasonable assumption, because the right-handed neutrinos N_{R} are $\text{SU}(2)_{\text{L}}$ singlets, they are irrelevant to the energy scale of the electroweak symmetry breaking, so we assume $M_{\mathrm{R}} \gg M_{\mathrm{D}}$. Multiplying Eq. (10.4.12) from left by $\left(\begin{smallmatrix} V & R \\ S & U \end{smallmatrix}\right)$ we obtain

$$\begin{cases} R\hat{M}_{\mathrm{N}} = M_{\mathrm{D}}U^* \\ S\hat{M}_{\nu} = M_{\mathrm{D}}^{\mathrm{T}}V^* + M_{\mathrm{R}}S^* \end{cases} \tag{10.4.18}$$

$$\begin{cases} U\hat{M}_{\mathrm{N}} = M_{\mathrm{D}}^{\mathrm{T}}R^* + M_{\mathrm{R}}U^* \\ \qquad V\hat{M}_{\nu} = M_{\mathrm{D}}S^* \end{cases} \tag{10.4.19}$$

When $M_{\mathrm{R}} \gg M_{\mathrm{D}}$, from Eqs. (10.4.18) and (10.4.19) we see that, $R \sim S \sim O(M_{\mathrm{D}}/M_{\mathrm{R}})$, U and V are almost unitary matrices with accuracy of $O(M_{\mathrm{D}}^2/M_{\mathrm{R}}^2)$ [28]. Eq. (10.4.19) leads to

$$U\hat{M}_{\mathrm{N}}U^{\mathrm{T}} = M_{\mathrm{D}}^{\mathrm{T}}(R^*U^{\mathrm{T}}) + M_{\mathrm{R}}(UU^+)^{\mathrm{T}} \approx M_{\mathrm{R}} \tag{10.4.20}$$

$$V\hat{M}_{\nu}V^{\mathrm{T}} = M_{\mathrm{D}}(S^*V^{\mathrm{T}}) \tag{10.4.21}$$

Multiplying the second equation of Eq. (10.4.18) from right-hand side by V^{T}, we obtain

$$S^*V^{\mathrm{T}} = M_{\mathrm{R}}^{-1}S\hat{M}_{\nu}V^{\mathrm{T}} - M_{\mathrm{R}}^{-1}M_{\mathrm{D}}^{\mathrm{T}}(VV^+)^{\mathrm{T}} \approx -M_{\mathrm{R}}^{-1}M_{\mathrm{D}}^{\mathrm{T}} \tag{10.4.22}$$

Substituting Eq. (10.4.22) into Eq. (10.4.21), we obtain

$$V\hat{M}_{\nu}V^{\mathrm{T}} \approx M_{\nu} \approx -M_{\mathrm{D}}M_{\mathrm{R}}^{-1}M_{\mathrm{D}}^{\mathrm{T}} \tag{10.4.23}$$

The first "≈" is for taking the matrices $R \sim S \sim 0$. This is consistent with the result of Eq. (10.4.7), i.e. the observed three types of neutrinos all get small masses.

For Type-II and Type-III See-saw mechanisms, please refer Ref. [27].

10.5. Parameterization of Neutrino Mixing Matrix

In Chapter 7, we have given two parameterization forms of Eqs. (7.7.5) and (7.7.6) for the quark mixing matrix V_{CKM}. For neutrino mixing matrix U_{MNS}, its form is similar with V_{CKM}, the only difference is that we need to include Majorana neutrinos. For detail, the readers can refer Ref. [20]. Here we only present the results below.

In general, U_{MNS} can be parameterized as

$$U_{\mathrm{MNS}} = (V_{\mathrm{CKM}}) \cdot P' \tag{10.5.1}$$

where

$$P' = \mathrm{diag}(\mathrm{e}^{i\rho}, \mathrm{e}^{\mathrm{i}\sigma}, 1) \tag{10.5.2}$$

is Majorana phase matrix. For quarks (d, s, b) or pure Dirac neutrinos, we can use field phase transformation to cancel the phases in P' (for example $d \to \mathrm{e}^{\mathrm{i}\varphi_d} d$, $s \to \mathrm{e}^{\mathrm{i}\varphi_s} s$, $b \to \mathrm{e}^{\mathrm{i}\varphi_b} b$), so we can make $P' = \mathrm{diag}(1, 1, 1)$, thus P' can be taken away, and only V_{CKM} is left.

But for Majorana neutrinos, because the restriction of the condition $\nu_i^C = \nu_i$, the transformations $\nu_i = \mathrm{e}^{\mathrm{i}\varphi_i} \nu_i$ make us to take only $\varphi_i = 0$, so the phases ρ, σ in P' cannot be canceled out.

10.6. *CP* Violation in Neutrino Oscillation

Neutrino oscillations can also be used to study CP, T, CPT violations.

- CP violation leads to $P(\nu_\alpha \to \nu_\beta; t) \neq P(\bar{\nu}_\alpha \to \bar{\nu}_\beta; t)$;
- T violation leads to $P(\nu_\alpha \to \nu_\beta; t) \neq P(\nu_\beta \to \nu_\alpha; t)$;
- CPT violation leads to $P(\nu_\alpha \to \nu_\beta; t) \neq P(\bar{\nu}_\beta \to \bar{\nu}_\alpha; t)$.

Below we only discuss the case of neutrino oscillation in vacuum under the condition of CPT invariance, because all discussions in this chapter are proceeded under the assumption of CPT invariance.

We first look at CP violation. According o Eqs. (10.2.17) and (10.2.20),

$$\begin{aligned}\Delta P_{CP} &= P(\nu_\alpha \to \nu_\beta) - P(\bar{\nu}_\alpha \to \bar{\nu}_\beta)\\ &= 16J \sum_\gamma \varepsilon_{\alpha\beta\gamma} \sin\left(\frac{\Delta m_{21}^2}{4E} L\right) \sin\left(\frac{\Delta m_{31}^2}{4E} L\right) \sin\left(\frac{\Delta m_{32}^2}{4E} L\right)\end{aligned} \tag{10.6.1}$$

From the definition of J in Eq. (7.10.14) we can see that Majorana phases ρ and σ do not go into the expression of J, i.e. J is irrelevant to ρ and σ, so ΔP_{CP} is also irrelevant to ρ and σ. Because

$$J = \cos\theta_{12} \cos^2\theta_{13} \cos\theta_{23} \sin\theta_{12} \sin\theta_{13} \sin\theta_{23} \sin\delta_{13} \tag{10.6.2}$$

Thus ΔP_{CP} is relative with the three mixing angles θ_{12}, θ_{13}, θ_{23}.

At present, PDG gives the world average values [29]

$$\begin{aligned}
\sin^2\theta_{12} &= 0.306^{+0.018}_{-0.015} \\
\sin^2\theta_{23} &= 0.42^{+0.08}_{-0.03} \\
\sin^2\theta_{13} &= 0.021^{+0.007}_{-0.008} \quad \text{or} \quad 0.0251 \pm 0.0034
\end{aligned} \tag{10.6.3}$$

At the reactor of Daya Bay, Guangdong, China, the $P(\bar{\nu}_e \to \bar{\nu}_e)$, i.e. the surviving probability of $\bar{\nu}_e$, was measured, so the decisive measurement of $\sin^2 2\theta_{13}$ was completed [30]. At the significance of 5.2σ the result is

$$\sin^2 2\theta_{13} = 0.092 \pm 0.016\,(\text{stat.}) \pm 0.005\,(\text{syst.}) \tag{10.6.4}$$

This leads to

$$\sin^2\theta_{13} \approx 0.0236 \tag{10.6.5}$$

It is consistent with world average value of Eq. (10.6.3). We should point out that Daya Bay Group is the earliest one which give the measurement of $\sin\theta_{13}$ with confident significance.

Because the CP violation ΔP_{CP} of Eq. (10.6.1) is too small, up to now there are no any experiments which observed neutrino CP violation effects.

For T violation, according to Eq. (10.2.17), the strength T violation

$$\begin{aligned}
\Delta P_T &= P(\nu_\alpha \to \nu_\beta) - P(\nu_\beta \to \nu_\alpha) \\
&= 16J \sum_\gamma \varepsilon_{\alpha\beta\gamma} \sin\left(\frac{\Delta m_{21}^2}{4E} L\right) \sin\left(\frac{\Delta m_{31}^2}{4E} L\right) = \Delta P_{CP}
\end{aligned} \tag{10.6.6}$$

This is obvious, $\Delta P_{CP} = \Delta P_T$ is just the result of the assumption of CPT invariance. Certainly, at present there are no any experiments which have observed T violation in neutrino oscillation.

10.7. Neutrino Mass Hierarchy Problem

Neutrino mass hierarchies are divided to two classes: normal and inverted. For normal hierarchy, $\Delta m_{32}^2 \sim \Delta m_{31}^2 \gg \Delta m_{21}^2 > 0$. For inverted, $\Delta m_{32}^2 \sim \Delta m_{31}^2 < 0$ (Fig. 10.7).

Because the matter density of the earth is not large, so for the oscillation experiments of the atmospheric neutrino ν_μ penetrating the earth and the accelerator produced neutrino ν_μ penetrating the earth (like MINOS) we all can ignore the matter effects. In these cases, the oscillation formulae

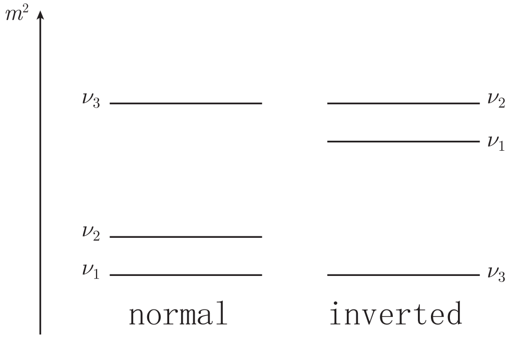


Fig. 10.7. Normal and inverted mass hierarchies.

(10.1.4) for only existing two flavors of neutrinos ν_μ and ν_τ is applicable. Obviously, from only Eq. (10.1.4) we cannot decide the sign of Δm^2. Therefore ν_μ oscillation experiments can only give $\left|\Delta m_{31}^2\right| \sim \left|\Delta m_{32}^2\right| \sim 2.4 \times 10^{-3}\,\mathrm{eV}^2$ (see Sec. 10.1).

The solar neutrino oscillation experiments are different. The matter density of the SUN is very large, the resonance condition in the process for solar neutrino (ν_e) penetrating the sun and arriving at the earth is (see Eqs. (10.2.38) and (10.2.39))

$$x = \frac{2\sqrt{2}G_\mathrm{F}N_\mathrm{e}E}{\Delta m^2} \approx \cos 2\theta \tag{10.7.1}$$

From Eq. (10.1.12), $\tan^2\theta = 0.47$, so we can get $\cos 2\theta > 0$. Thus Eq. (10.7.1) gives $x > 0$, this leads to

$$\Delta m^2 = \Delta m_{21}^2 > 0 \tag{10.7.2}$$

For either normal or inverted mass hierarchies, Eq. (10.1.12) is always correct, so

$$\Delta m_{21}^2 \sim 7.59 \times 10^{-5}\mathrm{eV}^2 > 0 \tag{10.7.3}$$

The signs of Δm_{32}^2 and Δm_{31}^2 are opposite for normal and inverted hierarchies, this will affect the measurement results for the physical quantities which are relevant to Δm_{32}^2, Δm_{31}^2. For example, T2K Group [31] used the accelerator produced ν_μ beam to measure $\nu_\mu \to \nu_\mathrm{e}$ oscillation probability

$P_{\mu\mathrm{e}} \equiv P(\nu_\mu \to \nu_\mathrm{e})$. Here the expression of $P_{\mu\mathrm{e}}$ is [32]

$$\begin{aligned} P_{\mu\mathrm{e}} \equiv P(\nu_\mu \to \nu_\mathrm{e}) &= \sin^2 2\theta_{13} \sin^2\theta_{23} \frac{\sin^2(A-1)\Delta}{(A-1)^2} \\ &\quad + \alpha \sin 2\theta_{13} \sin 2\theta_{12} \sin 2\theta_{23} \cos(\Delta + \delta) \frac{\sin A\Delta}{A} \frac{\sin(A-1)\Delta}{A-1} \\ &\quad + \alpha^2 \sin^2 2\theta_{12} \cos^2\theta_{23} \frac{\sin^2 A\Delta}{A^2} \end{aligned} \tag{10.7.4}$$

where

$$\Delta = \frac{\Delta m_{31}^2 L}{4E}, \quad A = \frac{2EV}{\Delta m_{31}^2}, \quad \alpha = \frac{\Delta m_{21}^2}{\Delta m_{31}^2}, \quad V = \sqrt{2} G_\mathrm{F} N_\mathrm{e} \tag{10.7.5}$$

δ is the CP phase in the mixing matrix. From Eqs. (10.7.4) and (10.7.5), we can see that the signs of Δm_{31}^2 and Δm_{32}^2 will affect the value of $P_{\mu\mathrm{e}}$, so are relevant to normal or inverted mass hierarchies. T2K measured $P_{\mu\mathrm{e}}$, and obtained the upper and lower limits of $\sin^2 2\theta_{13}$ [31]. For the case of CP phase $\delta = 0$, the normal (inverted) mass hierarchies are

$$0.03\ (0.04) < \sin^2 2\theta_{13} < 0.28\ (0.34) \tag{10.7.6}$$

The significance of the result is 2.5σ. The error is too large, so we cannot get any conclusion for the mass hierarchy just according to Eq. (10.110).

10.8. Perspectives

The problems for the future investigations are mainly the following aspects:

(1) *Signs of* Δm_{31}^2 *and* Δm_{32}^2: This problem has been already discussed in Sec. 10.7. It is actually the neutrino mass hierarch problem. At present, there are many groups which are trying to solve this problem. They will give the answer in the near future.

(2) *Are neutrinos dirac type or majorana type*: From mass eigenstate point of view, theoretically distinguishing the two types of neutrinos is very easy. Dirac neutrinos in Eq. (10.3.7) are different with their antiparticles (charge conjugate states), i.e. $\nu'^C \neq \nu'$. But for Majorana neutrinos in Eq. (10.3.16), their antiparticles are the same as themselves, i.e. $\nu'^C = \nu'$. Thus there is no problem for distinguishing them theoretically.

As for experiments, neutrinos participating weak interactions are not mass eigenstates but flavor eigenstates (interaction eigenstates). In that case, no matter Dirac or Majorana neutrinos, they all participate weak interactions through the flavor eigenstates $\nu_{\mathrm{L}} = (\nu_{\mathrm{eL}}, \nu_{\mu\mathrm{L}}, \nu_{\mu\mathrm{L}})$, so in general weak interaction processes cannot distinguish Dirac neutrinos from Majorana ones. But the mass term in Lagrangian of the Majorana neutrinos breaks lepton number conservation ($\Delta L = 2$) and induces $\nu - \bar{\nu}$ mixing, this mixing can create nuclear $0\nu 2\beta$ transition (see Fig. 10.6). Therefore, measuring nuclear $0\nu 2\beta$ transition can distinguish Majorana neutrinos from Dirac neutrinos. At present, nuclear $0\nu 2\beta$ transitions have not been found in all of the world laboratories.

(3) *Absolute measurements of neutrino masses*: In Sec. 10.4, we have given the cosmological observation constrain to neutrino masses $\sum_{i=\mathrm{e},\mu,\tau} m_i^2 < 0.28\mathrm{eV}$ [24], while the tritium β decay gives the recent most rigorous upper bound $m^2_{\bar{\nu}_{\mathrm{e}}} < 2.05\,\mathrm{eV}$ [25]. This kind of measurements will continue, and the mass upper bound will be deduced further. In addition, $0\nu 2\beta$ transition can also give constrain to neutrino masses.

(4) *CP violation search in neutrino oscillation*: As θ_{13} is very small, *CP* violation search can be proceeded in long baseline neutrino oscillations. In US, the accelerator PEP-II is rebuilt to fill the above possibility. Another method is to build neutrino factory. This needs to construct μ storage ring. We can get large amount of neutrinos from large amount of μ decays. But the biggest difficulty for constructing μ storage ring is how to collimate scattered μ particles coming from π decays into the storage ring, this is so-called μ cooling technology. This difficulty is under study in all of the world important laboratories.

(5) *Research on superhigh energy neutrino in Universe*: Recently, a new interdisciplinary study — neutrino astrophysics — is emerging. A cubic thousand meter scale of neutrino detector is under construction at the South Pole [33] (IceCube) and the cubic thousand meter scale of neutrino under water telescope is also under construction [34] (KM3NeT). These new detectors will play important role in the near future.

(6) *Detection of cosmic neutrino background (CνB)*: At present, there are no experimental results. Many experimental groups in the world are trying.

(7) *Research on neutrino cosmology*: Another new interdisciplinary study — neutrino cosmology — is also emerging. Scientists are investigating the role of neutrinos in the process of formation and development of the Universe. For example, is the creation of the baryon–antibaryon asymmetry relevant to neutrinos?

Certainly, many other problems need to be investigated.

References

1. Pauli W. *Proc. Solvary Congr.*, 1933: 324.
2. Danby G, *et al.* 1962 *Conference at CERN*, P809; Perl M L, *et al. Phys. Rev. Lett.*, 1975, 35: 1489.
3. Alpher R A, Bethe H, Gamo G W. *Phys. Rev.*, 1948, 73: 803.
4. Maki Z, Nakagawa M, Sakata S. *Prog. Theor. Phys.*, 1962, 28: 870.
5. Fukuda Y, *et al.* (Super-K Collab.). *Phys. Rev. Lett.*, 1998, 81: 1562.
6. Ashine Y, *et al.* (Super-K Collab.). *Phys. Rev.*, 2005, D71: 112005.
7. Ashine Y, *et al.* (Super-K Collab.). *Phys. Rev. Lett.*, 2004, 93: 101801.
8. Ahn M H, *et al.* (Super-K Collab.). *Phys. Rev. Lett.*, 2003, 90: 041801.
9. Michael D, *et al.* (MINOS Collab.). *Phys. Rev. Lett.*, 2006, 97: 191801.
10. Adamson P, *et al.* (MINOS Collab.). *Phys. Rev. Lett.*, 2008, 101: 131802.
11. Davis Jr. R, Harmer D S, Hoffman K C. *Phys. Rev. Lett.*, 1968, 20: 1205.
12. Hampel W, *et al. Phys. Lett.*, 1999, B447: 127; Altmann M, *et al. Phys. Lett.*, 2005, B616, 174; Abdurashitov J N, *et al. Phys. Rev.*, 2009, C80: 015087.
13. Bahcall J N, Pinsonneault M H, Basu S. *Astrophys. J*, 2001, 555: 990.
14. PDG Collab. *Phys. Rev.*, 2012, D86: 010001.
15. Ahmad Q R, *et al.* (SNO Collab.). *Phys. Rev. Lett.*, 2002, 89: 011301.
16. Aharmim B, *et al.* (SNO Collab.). *Phys. Rev.*, 2005, C72: 055502.
17. Bahcall J N, Serenelli A M, Basu S. *Astrophys. J*, 2005, 621: L85.
18. Eguchi K, *et al.* (KamLand Collab.). *Phys. Rev. Lett.*, 2003, 90: 021802; Abe S, *et al.* (KamL and Collab.). *Phys. Rev. Lett.*, 2008, 100: 221803.
19. An F P, *et al.* (Daya Bay Collab.). *Phys. Rev. Lett.*, 2012, 108: 171803.
20. Xing Z Z, Zhou S. *Neutinos in Particle Physics, Astronomy and Cosmology.* Zhejiang University Press, 2011.
21. Wolfenstein L. *Phys. Rev.*, 1978, D17: 2369; Mikheyev S, Smirnov A. *Sov. J. Nucl. Phys.*, 1985, 42: 913.
22. Kayser B. hep-ph, 2005, 0506165.
23. Majorana E. *Nuovo Cim*, 1937, 14: 171.
24. Thomas S A, Abdalla F B, Lahav O. *Phys. Rev. Lett.*, 2010, 105: 031301.
25. Aseev V N, *et al. Phys. Rev. D*, 2011, 84: 112003; Lobashev V, *et al. Nucl. Phys. A*, 2003, 719: 153c.
26. Fritzsch H, Gell-Mann M, Minkowski P. *Phys. Lett. B*, 1975, 59: 256; Minkowski P. *Phys. Lett. B*, 1977, 67: 421; Yanagida T. *Proc. Workshop on Unified Theory and the Baryon Number of the Universe.* KEK, 1979: 95; Gell-Mann M, Ramond P, Slansky R. *Supergravity.* North-Holland, 1979: 315; Glashow S L. *Quarks and Leptons.* Plenum, 1980: 707; Mohapatra R N, Senjanovic G. *Phys. Rev. Lett.*, 1980, 44: 912.
27. Konetschny W, Kummer W. *Phys. Lett. B*, 1977, 70: 443; Magg M, Wetterich C. *Phys. Lett. B*, 1980, 94: 61; Schechter J, Valle J W F. *Phys. Rev. D*, 1980, 22: 2227; Cheng T D, Li L F. *Phys. Rev. D*, 1980, 22: 2860; Lazarides G, Shafi Q, Wetterich C. *Nucl. Phys. B*, 1981, 181: 287; Mohapatra R N, Senjanovic G. *Phys. Rev. D*, 1981, 23: 265; Foot R, *et al. Z. Phys. C*, 1989, 44: 441; Ma E. *Phys. Rev. Lett.*, 1998, 81: 1171.
28. Xing Z Z, Zhou S. *High Ener. Phys. Nucl. Phys.* 2006, 30: 828.

29. Particle Data Group, *Phys. Rev. D*, 2012, 86: 010001.
30. An F P, *et al.* (Daya Bay Collab.). *Phys. Rev. Lett.*, 2012, 108: 171803.
31. Abe K, *et al.* (T2K Collab.). *Phys. Rev. Lett.*, 2011, 107: 041801.
32. Mezzetto M, Schwetz T, *J. Phys. G: Nucl. Part. Phys.*, 2010, 37: 103001.
33. Abbasi R, *et al.* (Ice Cube Collab.). *Astrophys. J.*, 2010, 710: 346.
34. Carr J, *et al.* (KM3NeT Collab.). 2007, 0711. 2145.

This page intentionally left blank

Chapter 11

Beyond Standard Model

The standard model here refers to QCD ($\mathrm{SU}(3)_\mathrm{C}$ gauge theory) for describing strong interactions and the electroweak unified theory ($\mathrm{SU}(2) \times \mathrm{U}(1)$ gauge theory). We call $\mathrm{SU}(3)_\mathrm{C} \times \mathrm{SU}(2) \times \mathrm{U}(1)$ gauge theory as standard model for describing strong, electromagnetic, and weak interactions. Till now, all experiments are consistent with the standard model predictions. But the standard model has its own shortcomings:

1. It is proved in neutrino oscillation experiments that neutrinos are massive (although tiny), but in the standard model neutrinos are exactly massless. For this point, the standard model should be at least extended simply.
2. In the standard model, CP violation attributes to the complex phase of the CKM quark mixing matrix, but the CP violating strength of this type is not big enough to create baryon–antibaryon asymmetry in the early evolution of our Universe, so we need to introduce new CP violation mechanisms.
3. The gauge group of the standard model is a direct product of three simple groups, thus has tree independent coupling constants g_3, g_2, g_1. So it is not real unification of all the interactions, even the electromagnetic and weak interactions are not really unified, although we still call $\mathrm{SU}(2) \times \mathrm{U}(1)$ gauge theory as the electroweak unified theory.
4. There is no explanation for the electric charge quantization.
5. In the standard model, there are no relation between quarks and leptons.
6. Why are there only three generations of quarks and leptons? Why the three generations fill up the same group representation repeatedly?

(7) There are too many free parameters in the standard model, at least 19. They are as follows:

- three gauge coupling constants g_3, g_2, g_1, or equivalently $\alpha_S = \frac{g_3^2}{4\pi}$, $\alpha_{\text{em}} = \frac{e^2}{4\pi}(e = g_2 \sin\theta_{\text{W}})$ and $\sin\theta_{\text{W}} = \frac{g_1^2}{(g_1^2+g_2^2)}$;
- nine fermion masses;
- four CKM mixing matrix parameters (3 mixing angles θ_{12}, θ_{23}, θ_{13}, 1 complex phase δ);
- two Higgs potential parameters, i.e. the vacuum expectation value υ and the ϕ^4 coefficient λ or equivalently m_Z, m_H;
- one strong CP violating parameter θ (or renormalized $\bar{\theta}$).
  If neutrino mixing matrix parameters are considered, there will be nine more new parameters, they are as follows:
- three neutrino masses m_{ν_1}, m_{ν_2}, m_{ν_3} (or equivalently $m_{\nu_{\text{e}}}$, m_{ν_μ}, m_{ν_τ});
- six parameters of lepton mixing matrix θ_1, θ_2, θ_3, δ, ρ, σ, the first four are similar with CKM matrix, the last two are from Majorana neutrino mixing matrix. Certainly here we assume there are three types of Majorana neutrinos.

(8) In the standard model, one loop Higgs mass correction encounters quadratic divergence, so the particle mass hierarchy in the standard model cannot be kept at loop diagram corrections.

All the above problems can only be solved beyond standard model. In what follows, we shall discuss some possibilities for extending the standard model.

11.1. Grand Unified Theories for Strong, Electromagnetic and Weak Interactions

11.1.1. *SU(5) grand unified theory*

One important direction beyond standard model is the grand unification of the strong, electromagnetic, and weak interactions. SU(5) grand unified theory (GUT) was proposed by Georgi and Glashow in 1974 [1]. Although SU(5) grand unified theory is ruled out by proton decay experiments (see discussion latter), but because of its simplicity and representative feature we still want to discuss it first and to explain the basic idea for grand unification. We should point out that super-symmetric SU(5) grand unified model is still not ruled out by proton decay experiments.

11.1.1.1. *Adler anomaly of SU(N) group representation*

For discussing the SU(5) model, the representations filled up by fermions must be guaranteed to have cancellation of Adler anomalies. Here we only give the results without derivation. For details, please refer Ref. [2].

The Adler anomalies, for $N \geq 3$, SU(N) group left-handed fermion representations, are as follows: for a totally antisymmetric representation described by an m-box column Young tableau, $\left.\begin{array}{|c|}\hline \ \\ \hline \ \\ \hline \vdots \\ \hline\end{array}\right\} m$ boxes, its Adler anomaly is

$$A_a\left(\begin{array}{|c|}\hline \ \\ \hline \ \\ \hline \vdots \\ \hline\end{array}\right) = \frac{(N-2m)(N-3)!}{(N-m-1)!(m-1)!} \tag{11.1.1}$$

For the representation described by m-box totally symmetric Young tableau, its anomaly is

$$A_s\left(\begin{array}{|c|c|c|}\hline \ & \ & \cdots \\ \hline\end{array}\right) = \frac{(N+2m)(N+m)}{(N+2)!(m-1)!} \tag{11.1.2}$$

From Eqs. (11.1.1) and (11.1.2), we obtain

$$A(\square)=1,\ A\left(\begin{array}{|c|}\hline \ \\ \hline \ \\ \hline\end{array}\right)=N-4,\ A\left(\begin{array}{|c|}\hline \ \\ \hline \ \\ \hline \ \\ \hline\end{array}\right)=\frac{1}{2}(N-3)(N-6),\ A\left(\begin{array}{|c|c|}\hline \ & \ \\ \hline\end{array}\right)=N+4 \tag{11.1.3}$$

etc.

For right-handed fermion representations, their Adler anomalies change sign, that is, for representation R,

$$A(R_R) = -A(R_L) \tag{11.1.4}$$

For the complex conjugate representation R^*, we have

$$A(R^*) = -A(R) \tag{11.1.5}$$

For SU(5) group, because $\begin{array}{|c|}\hline \ \\ \hline \ \\ \hline \ \\ \hline\end{array} = \begin{array}{|c|}\hline \ \\ \hline \ \\ \hline\end{array}^{\,*}$, we have

$$A\left(\begin{array}{|c|}\hline \ \\ \hline \ \\ \hline \ \\ \hline\end{array}\right) = -A\left(\begin{array}{|c|}\hline \ \\ \hline \ \\ \hline\end{array}\right) \tag{11.1.6}$$

For the product and summation of representations, Adler anomalies satisfy

$$A(R_1 \times R_2) = D(R_1)A(R_2) + A(R_1)D(R_2) \tag{11.1.7}$$

$$A(R_1 + R_2) = A(R_1) + A(R_2) \tag{11.1.8}$$

where $D(R)$ is the dimension of the representation R, and $A(R)$ is its Adler anomaly.

In Sec. 5.1, we already give the decompositions of Young tableau products and the formulas for calculating the corresponding representation dimensions. From Eqs. (11.1.7) and (11.1.8) and decomposition of representations by use of Young Tableau, we can compute out the Adler anomalies corresponding to any Young tableau.

11.1.1.2. *Filling representations for fermions and gauge bosons in the SU(5) grand unified model*

For describing a massive fermion field ψ, we need two Weyl spinors

$$\psi_{\mathrm{L}} = \frac{1-\gamma_5}{2}\psi, \quad \psi_{\mathrm{R}} = \frac{1+\gamma_5}{2}\psi \tag{11.1.9}$$

Here, we use Bjorken–Drell matrix and γ matrix definition. Define

$$\left.\begin{aligned} \psi^C = C\bar{\psi}^{\mathrm{T}}, \psi_{\mathrm{L}}^C = \left(\psi^C\right)_{\mathrm{L}} &= \frac{1-\gamma_5}{2}\psi^C \\ \psi_{\mathrm{R}}^C = \left(\psi^C\right)_{\mathrm{R}} &= \frac{1+\gamma_5}{2}\psi^C \end{aligned}\right\} \tag{11.1.10}$$

So a massive fermion can be described by a four-component Dirac spinor ψ or equivalently any one of the following three pairs of spinors, i.e.

$$(\psi_{\mathrm{L}}, \psi_{\mathrm{R}}), \quad (\psi_{\mathrm{L}}, \psi_{\mathrm{L}}^C), \quad (\psi_{\mathrm{R}}, \psi_{\mathrm{R}}^C) \tag{11.1.11}$$

In the standard model, the first generation of fermions includes u_α, d_α, e, ν_{e} ($\alpha = 1, 2, 3$ are color indices). There are totally 15 Weyl spinors, i.e.

$$\begin{pmatrix} \mathrm{u}_{\mathrm{L}}^\alpha \\ \mathrm{d}_{\mathrm{L}}^\alpha \end{pmatrix}, \quad \mathrm{u}_{\mathrm{R}}^\alpha, \quad \mathrm{d}_{\mathrm{R}}^\alpha, \quad \begin{pmatrix} \nu_{\mathrm{eL}} \\ \mathrm{e}_{\mathrm{eL}} \end{pmatrix}, \quad \mathrm{e}_{\mathrm{R}} \tag{11.1.12}$$

These 15 components should fill up a representation of SU(5) with Adler anomaly cancellation. One possible choice can be

$$\begin{array}{|c|}\hline \\ \hline \\ \hline \end{array}_{\,\mathrm{L}} + \begin{array}{|c|}\hline \\ \hline \\ \hline \\ \hline \\ \hline \end{array}_{\,\mathrm{L}} = 10_{\mathrm{L}} + 5_{\mathrm{L}}^* \tag{11.1.13}$$

or equivalently

$$\begin{array}{|c|}\hline \\ \hline \\ \hline \end{array}_{\mathrm{L}} + \square_{\mathrm{R}} = 10_{\mathrm{L}} + 5_{\mathrm{R}} \tag{11.1.14}$$

From Eqs. (11.1.1), (11.1.2) and (11.1.5), we see that Eqs. (11.1.13) and (11.1.14) are all good choice. In the following, we use Eq. (11.1.14), that is, we put

$$\psi_{\mathrm{R}}^{i} \equiv 5_{\mathrm{R}}, \quad \psi_{\mathrm{L}}^{ij} \equiv 10_{\mathrm{L}} \tag{11.1.15}$$

ψ_{L}^{ij} is antisymmetric with respect to i, j, $i, j = 1, 2, \ldots, 5$. The detailed form is

$$\psi_{\mathrm{R}}^{i} \equiv 5_{\mathrm{R}} = \begin{pmatrix} \mathrm{d}_1 \\ \mathrm{d}_2 \\ \mathrm{d}_3 \\ \mathrm{e}^c \\ -\nu_{\mathrm{e}}^c \end{pmatrix}_{\mathrm{R}} \tag{11.1.16}$$

where $\begin{pmatrix} \mathrm{e}^c \\ -\nu_{\mathrm{e}}^c \end{pmatrix}$ comes from $\begin{pmatrix} \nu_{\mathrm{e}} \\ \mathrm{e} \end{pmatrix} \to \mathrm{i}\sigma_2 \begin{pmatrix} \nu_{\mathrm{e}} \\ \mathrm{e} \end{pmatrix}^* = \begin{pmatrix} \mathrm{e}^c \\ -\nu_{\mathrm{e}}^c \end{pmatrix}$. Moreover,

$$\psi_{\mathrm{L}}^{ij} \equiv 10_{\mathrm{L}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & \mathrm{u}_3^c & -\mathrm{u}_2^c & -\mathrm{u}_1 & -\mathrm{d}_1 \\ -\mathrm{u}_3^c & 0 & \mathrm{u}_1^c & -\mathrm{u}_2 & -\mathrm{d}_2 \\ \mathrm{u}_2^c & -\mathrm{u}_1^c & 0 & -\mathrm{u}_3 & -\mathrm{d}_3 \\ \mathrm{u}_1 & \mathrm{u}_2 & \mathrm{u}_3 & 0 & -\mathrm{e}^c \\ \mathrm{d}_1 & \mathrm{d}_2 & \mathrm{d}_3 & \mathrm{e}^c & 0 \end{pmatrix}_{\mathrm{L}} \tag{11.1.17}$$

From Eq. (11.1.16), we see that the electric charge operator Q is diagonal

$$Q = \begin{pmatrix} -1/3 & & & & \\ & -1/3 & & & \\ & & -1/3 & & \\ & & & 1 & \\ & & & & 0 \end{pmatrix} \tag{11.1.18}$$

For the filling of gauge fields, because SU(5) has 24 independent generators, there are 24 gauge fields $A_{\mu}^{a}(x)(a = 1, 2, \ldots, 24)$. For writing down the specific expressions, we firstly give the SU(5) generators.

The SU(5)group elements are as follows:

$$\begin{cases} u = \mathrm{e}^{\mathrm{i}\frac{1}{2}T_a\theta_a} \\ \mathrm{Tr}\ T_a^2 = 2, \quad a = 1, 2, \ldots, 24 \end{cases} \tag{11.1.19}$$

where

$$T_1 = \begin{pmatrix} \lambda_1 & 0 & \\ 0 & 0 & 0 \\ & 0 & 0 \end{pmatrix}, \quad T_2 = \begin{pmatrix} \lambda_2 & 0 & \\ 0 & 0 & 0 \\ & 0 & 0 \end{pmatrix}, \ldots, \quad T_8 = \begin{pmatrix} \lambda_8 & 0 & \\ 0 & 0 & 0 \\ & 0 & 0 \end{pmatrix}$$

$$T_9 = \begin{pmatrix} & & & 1 & 0 \\ & 0 & & 0 & 0 \\ & & & 0 & 0 \\ 1 & 0 & 0 & & \\ 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & & \end{pmatrix}, \quad T_{10} = \begin{pmatrix} & & & -\mathrm{i} & 0 \\ & 0 & & 0 & 0 \\ & & & 0 & 0 \\ \mathrm{i} & 0 & 0 & & \\ 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & & \end{pmatrix}$$

$$T_{11} = \begin{pmatrix} & & & 0 & 0 \\ & 0 & & 1 & 0 \\ & & & 0 & 0 \\ 0 & 1 & 0 & & \\ 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & & \end{pmatrix},$$

$$T_{12} = \begin{pmatrix} & & & 0 & 0 \\ & 0 & & -i & 0 \\ & & & 0 & 0 \\ 0 & i & 0 & & \\ 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & & \end{pmatrix}, \ldots \quad T_{15} = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & -3 & \\ & & & & 0 \end{pmatrix},$$

$$T_{16} = \begin{pmatrix} & & & 0 & 1 \\ & 0 & & 0 & 0 \\ & & & 0 & 0 \\ 0 & 0 & 0 & & \\ 1 & 0 & 0 & 0 & \\ 0 & 0 & 0 & & \end{pmatrix}, \ldots, \quad T_{24} = \frac{1}{\sqrt{10}} \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & -4 \end{pmatrix} \tag{11.1.20}$$

where $\lambda_i (i = 1, 2, \ldots, 8)$ are Gell-Mann SU(3) group generators

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_2 = \begin{pmatrix} 0 & -\mathrm{i} & 0 \\ \mathrm{i} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \ldots, \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & & \\ & 1 & \\ & & -2 \end{pmatrix} \tag{11.1.21}$$

Define

$$A_\mu = \frac{1}{2} \sum_a T_a A_{a\mu} = \frac{1}{\sqrt{2}} \hat{A}_\mu \tag{11.1.22}$$

Here only $\hat{A}_\mu$ can have correct normalization, this point can be seen clearly in Eq. (11.1.25). Using the expressions of T_a in Eq. (11.1.20), we have

$$\begin{aligned} A &= \frac{1}{\sqrt{2}} \hat{A} \\ &= \begin{pmatrix} \frac{A_3}{2} + \frac{A_8}{2\sqrt{3}} + \frac{A_{15}}{2\sqrt{6}} + \frac{A_{24}}{2\sqrt{10}} & \frac{A_1 - \mathrm{i}A_2}{2} & \frac{A_4 - \mathrm{i}A_5}{2} & \frac{A_9 - \mathrm{i}A_{10}}{2} & \frac{A_{16} - \mathrm{i}A_{17}}{2} \\ \frac{A_1 + \mathrm{i}A_2}{2} & \frac{-A_3}{2} + \frac{A_8}{2\sqrt{3}} + \frac{A_{15}}{2\sqrt{6}} + \frac{A_{24}}{2\sqrt{10}} & \frac{A_6 - \mathrm{i}A_7}{2} & \frac{A_{11} - \mathrm{i}A_{12}}{2} & \frac{A_{18} - \mathrm{i}A_{19}}{2} \\ \frac{A_4 + \mathrm{i}A_5}{2} & \frac{A_6 + \mathrm{i}A_7}{2} & \frac{-2A_8}{2\sqrt{3}} + \frac{A_{15}}{2\sqrt{6}} + \frac{A_{24}}{2\sqrt{10}} & \frac{A_{13} - \mathrm{i}A_{14}}{2} & \frac{A_{20} - \mathrm{i}A_{21}}{2} \\ \frac{A_9 + \mathrm{i}A_{10}}{2} & \frac{A_{11} + \mathrm{i}A_{12}}{2} & \frac{A_{13} + \mathrm{i}A_{14}}{2} & \frac{-3A_{15}}{2\sqrt{6}} + \frac{A_{24}}{2\sqrt{10}} & \frac{A_{22} - \mathrm{i}A_{23}}{2} \\ \frac{A_{16} + \mathrm{i}A_{17}}{2} & \frac{A_{18} + \mathrm{i}A_{19}}{2} & \frac{A_{20} + \mathrm{i}A_{21}}{2} & \frac{A_{22} + \mathrm{i}A_{23}}{2} & \frac{-4A_{24}}{2\sqrt{10}} \end{pmatrix} \end{aligned} \tag{11.1.23}$$

Define

$$
\begin{aligned}
&\frac{1}{\sqrt{2}}\left(\frac{A_{15}}{\sqrt{6}}+\frac{A_{24}}{\sqrt{10}}\right)=\frac{-2B}{\sqrt{30}}\\
&\frac{1}{\sqrt{2}}\left(\frac{-3A_{15}}{\sqrt{6}}+\frac{A_{24}}{\sqrt{10}}\right)=\frac{W^3}{\sqrt{2}}+\frac{3B}{\sqrt{30}}\\
&\frac{-4}{\sqrt{2}}\cdot\frac{A_{24}}{\sqrt{10}}=-\frac{W^3}{\sqrt{2}}+\frac{3B}{\sqrt{30}}\\
&W^+=\frac{A_{22}-\mathrm{i}A_{23}}{\sqrt{2}},\quad W^-=\frac{A_{22}+\mathrm{i}A_{23}}{2}\\
&G_2^1=\frac{A_1-\mathrm{i}A_2}{\sqrt{2}},\quad G_3^1=\frac{A_4-\mathrm{i}A_5}{\sqrt{2}},\quad G_3^2=\frac{A_6-\mathrm{i}A_7}{\sqrt{2}}\\
&G_1^1=\frac{A_3}{\sqrt{2}}+\frac{A_8}{\sqrt{6}},\quad G_2^2=\frac{-A_3}{\sqrt{2}}+\frac{A_8}{\sqrt{6}},\quad G_3^3=-\frac{2A_8}{\sqrt{6}}\\
&X_1=\frac{A_9+\mathrm{i}A_{10}}{\sqrt{2}},\quad X_2=\frac{A_{11}+\mathrm{i}A_{12}}{\sqrt{2}},\quad X_3=\frac{A_{13}+\mathrm{i}A_{14}}{\sqrt{2}}
\end{aligned}
\tag{11.1.24}
$$

$\bar{X}_j (j=1,2,3)$ can be obtained by changing $\mathrm{i}\to-\mathrm{i}$ in the expressions of X_j in Eq. (11.1.24). While

$$
Y_1=\frac{A_{16}+\mathrm{i}A_{17}}{\sqrt{2}},\quad Y_2=\frac{A_{18}+\mathrm{i}A_{19}}{\sqrt{2}},\quad Y_3=\frac{A_{20}+\mathrm{i}A_{21}}{\sqrt{2}}
$$

Again $\bar{Y}_j (j=1,2,3)$ can be obtained by changing $\mathrm{i}\to-\mathrm{i}$ in the above Y_j expressions.

Therefore Eq. (11.1.23) can be rewritten as

$$
\hat{A}=\begin{pmatrix}
G_1^1-\frac{2B}{\sqrt{30}} & G_2^1 & G_3^1 & \bar{X}_1 & \bar{Y}_1\\
G_1^2 & G_2^2-\frac{2B}{\sqrt{30}} & G_3^2 & \bar{X}_2 & \bar{Y}_2\\
G_1^3 & G_2^3 & G_3^3-\frac{2B}{\sqrt{30}} & \bar{X}_3 & \bar{Y}_3\\
X_1 & X_2 & X_3 & \frac{W^3}{\sqrt{2}}+\frac{3B}{\sqrt{30}} & W^+\\
Y_1 & Y_2 & Y_3 & W^- & \frac{-W^3}{\sqrt{2}}+\frac{3B}{\sqrt{30}}
\end{pmatrix}
\tag{11.1.25}
$$

Now we already fill in the first generation of fermions and the 24 gauge fields. $G^i_j(i,j=1,2,3)$ have only eight independent gluon fields. In the SU(5) grand unified model, in order to accommodate three generations of quarks and leptons we need repeat $\underline{10}_{\mathrm{L}}+\underline{5}_{\mathrm{R}}$ representation three times, this is the shortcoming of the model.

11.1.1.3. *Higgs mechanism*

The SU(5) symmetry must be broken down to $\mathrm{SU}(3)_{\mathrm{C}}\times\mathrm{SU}(2)\times\mathrm{U}(1)$ of the standard model at some very high-energy scale. This energy scale should be decided by the intersection point of the three coupling constants of $\mathrm{SU}(3)_{\mathrm{C}}$, SU(2), U(1) when the momentum transfer Q^2 increases gradually. This is depicted in Fig. 11.1. The revolutions of all the coupling constants $\alpha_i(Q^2)(i=1,2,3)$ along with Q^2 are decided by renormalization group equations. The results show that $\alpha_3(Q^2)$, $\alpha_2(Q^2)$, $\alpha_1(Q^2)$ basically intersect at $Q\approx10^{15}$GeV, where $\alpha_i(Q^2)=\frac{g_i^2(Q^2)}{4\pi}$.

According to physics requirements, the $\mathrm{SU}(3)_{\mathrm{C}}\times\mathrm{SU}(2)\times\mathrm{U}(1)$ gauge symmetry should also be broken at the scale of $Q\sim M_Z\sim100\,\mathrm{GeV}$ further into $\mathrm{SU}(3)_{\mathrm{C}}\times\mathrm{U}(1)_{\mathrm{em}}$. So the spontaneous symmetry breaking chain of the SU(5) grand unified model is

$$\mathrm{SU}(5)\xrightarrow{M_X\sim10^{15}\ \mathrm{GeV}}\mathrm{SU}(3)_{\mathrm{C}}\times\mathrm{SU}(2)\times\mathrm{U}(1)\xrightarrow{M_Z\sim100\,\mathrm{GeV}}\mathrm{SU}(3)_{\mathrm{C}}\times\mathrm{U}(1)_{\mathrm{em}} \tag{11.1.26}$$

The first step of breaking can be achieved by introducing 24-dimensional adjoint Higgs fields $\Phi=(\phi^i_j)(\phi^i_i=0)$. Here the repeated indices mean to

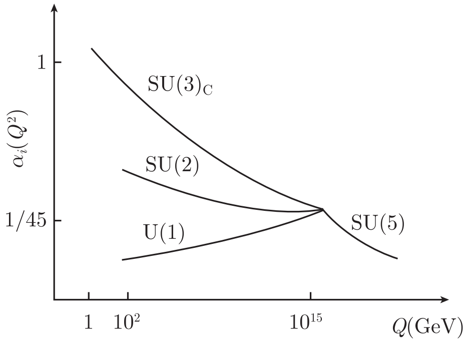


Fig. 11.1. Evolution of the coupling constants of the SU(5) grand unified model along with energy scale.

be summed over. The vacuum expectation value of ϕ^i_j is a diagonal matrix (here Φ is 5×5 matrix),

$$\left\langle \phi^i_j \right\rangle = \begin{pmatrix} -2/3 & & & & \\ & -2/3 & & & \\ & & -2/3 & & \\ & & & 1 & \\ & & & & 1 \end{pmatrix} a \tag{11.1.27}$$

where $a \sim 10^{15}$ GeV. It is easy to prove that the vacuum expectation value in Eq. (11.1.27) breaks SU(5) symmetry, but keeps $\mathrm{SU}(3)_\mathrm{C} \times \mathrm{SU}(2) \times \mathrm{U}(1)$ symmetry unchanged.

In fact, put u as the transformation matrix of the subgroup $\mathrm{SU}(3)_\mathrm{C}$, then ϕ^i_j transform under $\mathrm{SU}(3)_\mathrm{C}$ into

$$\phi^i_j \to \phi^{i'}_j = (u^+)^k_j u^i_l \phi^l_k, \quad i, j, k, l = 1, 2, 3 \tag{11.1.28}$$

the vacuum expectation value becomes

$$\left\langle \phi^i_j \right\rangle \to \langle \phi^{i'}_j \rangle = (u^+)^k_j u^i_l \left\langle \phi^l_k \right\rangle \tag{11.1.29}$$

From Eq. (11.1.27), for $i, j = 1, 2, 3$

$$\left\langle \phi^i_j \right\rangle = -\frac{2}{3} a \delta^i_j \tag{11.1.30}$$

Equation (11.1.28) changes to

$$\begin{aligned} \langle \phi^{i'}_j \rangle &= (u^+)_{j^k} u^i_l (-\frac{2}{3} a) \delta^k_l = -\frac{2}{3} a (u^+)^k_j u^i_k \\ &= -\frac{2}{3} a \delta^i_J = \left\langle \phi^i_j \right\rangle \end{aligned} \tag{11.1.31}$$

In the last step, $u^+ u = 1$ has been used. Equation (11.1.31) shows that the expectation value $\left\langle \phi^i_j \right\rangle$ is invariant under $\mathrm{SU}(3)_\mathrm{C}$ transformation. Similarly, for the SU(2) subgroup, taking $i, j, k, l = 4, 5$ in the above formulas we can easily prove that $\left\langle \phi^i_j \right\rangle$ is also invariant under transformations of the SU(2) subgroup.

Now let us find the remaining U(1) subgroup. Assume that the generator of this remaining U(1) group is Y and $\mathrm{Tr}\, Y^2 = 2$, then when $\mathrm{SU}(5) \to \mathrm{SU}(3)_\mathrm{C} \times \mathrm{SU}(2) \times \mathrm{U}(1)_Y$, Y should be orthogonal to all the generators of $\mathrm{SU}(3)_\mathrm{C}$ and SU(2).

The diagonal generators of $\mathrm{SU}(3)_{\mathrm{C}}$ are as follows:

$$T_3 = \begin{pmatrix} \lambda_3 & 0 \\ 0 & 0 \end{pmatrix}, \quad T_8 = \begin{pmatrix} \lambda_8 & 0 \\ 0 & 0 \end{pmatrix} \tag{11.1.32}$$

The diagonal generator of SU(2) is as follows:

$$T_3' = \begin{pmatrix} 0 & 0 \\ 0 & \sigma_3 \end{pmatrix} \tag{11.1.33}$$

where λ_3, λ_8 are SU(3) Gell-Mann matrices in Eq. (11.1.21), σ_3 is Pauli matrix $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. According to

$$\begin{cases} \mathrm{Tr}(YT_3) = 0 \\ \mathrm{Tr}(YT_8) = 0 \\ \mathrm{Tr}(YT_3') = 0 \\ \mathrm{Tr}\,Y = 0 \\ \mathrm{Tr}\,Y^2 = 2 \end{cases} \tag{11.1.34}$$

we can decide Y uniquely as

$$Y = \sqrt{\frac{5}{3}} \begin{pmatrix} -2/3 & & & & \\ & -2/3 & & & \\ & & -2/3 & & \\ & & & 1 & \\ & & & & 1 \end{pmatrix} \tag{11.1.35}$$

Now let us prove that $\langle \phi^i_j \rangle$ is invariant under $\mathrm{U}(1)_Y$ transformation. For this purpose, we only need to prove that $\left\langle \phi^i_j \right\rangle$ is invariant under the infinitesimal transformation $\mathrm{U}(1)_Y \sim 1 + \mathrm{i}Y\theta/2$. From Eq. (11.1.27), $\left\langle \phi^l_k \right\rangle = \omega_k \delta^l_k$, where for $k = 1, 2, 3,\ \omega_k = -\frac{2}{3}a$; for $k = 4, 5, \omega_k = a$. Substituting the infinitesimal transformation of $\mathrm{U}(1)_Y = \mathrm{e}^{\mathrm{i}\frac{\theta}{2}Y}$ into Eq. (11.1.27)

$$\begin{aligned} \left\langle \phi^i_j \right\rangle \to \langle \phi^{i'}_j \rangle &= \left(1 - \frac{\mathrm{i}}{2} Y\theta\right)^k_j \left(1 + \mathrm{i}\frac{1}{2} Y\theta\right)^i_l \omega_k \delta^l_k \\ &= \omega_k \left(1 - \mathrm{i}\frac{Y}{2}\theta\right)^k_j \left(1 + \mathrm{i}\frac{Y}{2}\theta\right)^i_k \\ &= \omega_i \delta^i_j - \frac{\mathrm{i}}{2}\theta(\omega_i - \omega_j) \mathrm{Y}^i_j \\ &= \omega_i \delta^i_j = \left\langle \phi^i_j \right\rangle \end{aligned} \tag{11.1.36}$$

This prove that $\left\langle\phi_j^i\right\rangle$ breaks SU(5) indeed down to $\mathrm{SU}(3)_{\mathrm{C}} \times \mathrm{SU}(2) \times \mathrm{U}(1)_Y$.

The second step breaking is completed by use of Higgs fundamental representation φ^i, where the vacuum expectation value of φ^i takes the following form:

$$\left\langle\varphi^i\right\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \upsilon \end{pmatrix}, \quad \upsilon \sim M_Z \sim 100\,\mathrm{GeV} \tag{11.1.37}$$

For any transformation

$$u = \mathrm{e}^{\mathrm{i}\frac{1}{2}T_a\theta_a} = \sum_{n=0}^{\infty}\left(\frac{\mathrm{i}}{2}\right)^n \frac{\theta_a^n T_a^n}{n!} \tag{11.1.38}$$

For

$$u \in \mathrm{SU}(3)_C,$$
$$u = \begin{pmatrix} 3\times 3 & & 0 \\ & 0 & 0 \\ 0 & & \\ & 0 & 0 \end{pmatrix},$$

the generator of $\mathrm{SU}(3)_{\mathrm{C}}$ T_a takes the same form, therefore

$$T_a\left\langle\varphi^i\right\rangle = \begin{pmatrix} 3\times 3 & & 0 \\ & 0 & 0 \\ 0 & & \\ & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \upsilon \end{pmatrix} = 0 \tag{11.1.39}$$

So $u\left\langle\varphi^i\right\rangle = \left\langle\varphi^i\right\rangle$, i.e. $\mathrm{SU}(3)_{\mathrm{C}}$ keeps invariant under the second step breaking. For $u \in \mathrm{SU}(2)$, the generator $T_a = \left(\begin{smallmatrix} 0 & 0 \\ 0 & \sigma_i \end{smallmatrix}\right)$, $i = 1, 2, 3$. Obviously

$$T_a\left\langle\varphi^i\right\rangle \neq 0 \tag{11.1.40}$$

so $u\left\langle\varphi^i\right\rangle \neq \left\langle\varphi^i\right\rangle$, SU(2) symmetry is broken.

For $u \in \mathrm{U}(1)$, $Y\left\langle\varphi^i\right\rangle \neq 0$, this leads to $u\left\langle\varphi^i\right\rangle \neq \left\langle\varphi^i\right\rangle$, so $\mathrm{U}(1)_Y$ symmetry is also broken. But from Eq. (11.1.18) we see that

$$Q\left\langle\varphi^i\right\rangle = 0 \tag{11.1.41}$$

It shows that although $\mathrm{U}(1)_Y$ symmetry is broken, but the $\mathrm{U}(1)_{\mathrm{em}}$ symmetry with generator Q is still kept. Finally

$$\mathrm{SU}(3)_{\mathrm{C}} \times \mathrm{SU}(2) \times \mathrm{U}(1)_Y \xrightarrow{\langle\varphi^i\rangle} \mathrm{SU}(3)_{\mathrm{C}} \times \mathrm{U}(1)_{\mathrm{em}} \tag{11.1.42}$$

11.1.1.4. *Fermion masses*

There are two types of Yukawa couplings in SU(5) grand unified model, i.e.

$$L_Y = A\bar{\psi}_{ij\mathrm{L}}\varphi^i\psi_{\mathrm{R}}^j + \text{h.c.} \tag{11.1.43}$$

$$L'_Y = B\frac{2}{2!2!}\varepsilon_{ijklm}\psi_{\mathrm{L}}^{T_{ij}}C^{-1}\varphi^k\psi_{\mathrm{L}}^{lm} + \text{h.c.} \tag{11.1.44}$$

where A, B are coupling constants, C is Dirac charge conjugate matrix, ε_{ijklm} is Levy-Civita tensor, $i, j, h, l, m = 1, 2, \ldots, 5$. Equations (11.1.43) and (11.1.44) are all SU(5) invariant couplings. We cannot construct an SU(5) invariant Yukawa coupling using ψ_{L}^{ij}, ψ_{R}^k and ϕ_j^i, so in the SU(5) grand unified model we cannot have super-heavy fermions, i.e. we cannot have fermions with mass at the order of $\langle\phi_j^i\rangle$.

Substituting $\langle\varphi^i\rangle$, $\langle\phi_j^i\rangle$ into Eqs. (11.1.43) and (11.1.44) we have

$$L_Y = Av(\bar{d}_{\mathrm{L}}d_{\mathrm{R}} + \bar{e}_{\mathrm{L}}e_{\mathrm{R}}) + \text{h.c.} \tag{11.1.45}$$

$$L'_Y = 2Bv(\bar{u}_{\mathrm{L}}u_{\mathrm{R}}) + \text{h.c.} \tag{11.1.46}$$

From Eqs. (11.1.45) and (11.1.46), we obtain

$$\begin{cases} m_{\mathrm{d}} = m_{\mathrm{e}} = Av \\ m_{\mathrm{u}} = 2Bv \\ m_{\nu_{\mathrm{e}}} = 0 \end{cases} \tag{11.1.47}$$

Equation (11.1.47) is only applicable at the grand unification energy scale $M_{\mathrm{X}} \sim 10^{15}\,\mathrm{GeV}$. The reason is that the coupling constants A, B take the same value for all fermions only at SU(5) invariant couplings, i.e. the grand unification scale. After SU(5) symmetry breaking, A, B will change along with the variation of the energy scale. Therefore, after the breakdown of SU(5) symmetry, A, B will be different for quarks and leptons. Specially, the effects of the evolution of renormalization group equation with energy scale are stronger for hadrons (quarks) than for leptons (leptons have only electroweak interactions), so for three generations of quarks and leptons at the GUT scale

$$m_{\mathrm{d}} = m_{\mathrm{e}}, \quad m_{\mathrm{s}} = m_{\mu}, \quad m_{\mathrm{b}} = m_{\tau} \tag{11.1.48}$$

Using renormalization equations, we can compute out [3]

$$m_{\mathrm{b}}\,(10\,\mathrm{GeV}) \sim 3m_{\tau} \sim 5\,\mathrm{GeV} \tag{11.1.49}$$

this is compatible with experiments. But for light quarks, form Eq. (11.1.48)

$$m_{\mathrm{d}}/m_{\mathrm{s}} = m_{\mathrm{e}}/m_{\mu} \sim 1/200 \tag{11.1.50}$$

This ratio is irrelevant with renormalization group, thus contradictory with the experimental value of $m_{\mathrm{d}}/m_{\mathrm{s}} \sim 1/20$.

In the SU(5) grand unified model discussed above, there are only 24-dimensional adjoint representation Higgs ϕ^i_j and 5-dimensional fundamental representation Higgs φ^i. This model is called minimal SU(5) grand unified model. If we want to solve the problem of Eq. (11.1.50), we can introduce three 5-dimensional representation Higgs $\varphi^i_\alpha(\alpha = 1, 2, 3)$ and a 45-dimensional representation Higgs ϕ^i_{jk} [4]. In that case at the GUT energy scale we can have the following mass relations

$$\begin{cases} m_{\mathrm{b}} = m_{\tau} \\ m_{\mathrm{d}} = 3m_{\mathrm{e}} \\ m_{\mathrm{s}} = \dfrac{1}{3}m_{\mu} \end{cases} \tag{11.1.51}$$

so we can have

$$m_{\mathrm{d}}/m_{\mathrm{s}} = 9m_{\mathrm{e}}/m_{\mu} \sim 1/20 \tag{11.1.52}$$

Obviously, from Eq. (11.1.47) the mass of u quark (c, t quarks) is a free parameter; there is no theoretical prediction. For neutrinos because B–L is still conserved, the minimal SU(5) GUT model cannot accommodate Dirac or Majorana mass term, so neutrinos are massless accurately. We shall not discuss this problem here.

11.1.1.5. *Masses of gauge bosons*

The heavy gauge bosons at GUT energy scale get masses at the first step symmetry breaking, i.e. $\mathrm{SU}(5) \xrightarrow{10^{15}\,\mathrm{GeV}} \mathrm{SU}(3)_{\mathrm{C}} \times \mathrm{SU}(2) \times \mathrm{U}(1)$. The quadratic term of the gauge bosons comes from

$$(D^{\mu}\Phi)^{+}(D_{\mu}\Phi) \tag{11.1.53}$$

where,

$$\begin{cases} D_{\mu}\Phi = \partial_{\mu}\Phi - \mathrm{i}g[A_{\mu},\Phi] \\ (D^{\mu}\Phi)^{+} = \partial_{\mu}\Phi^{+} + \mathrm{i}g[A^{\mu},\Phi]^{+} \\ A_{\mu} = \dfrac{1}{2}T_{a}A_{\mu}^{a} \end{cases} \tag{11.1.54}$$

In the matrix equation (11.1.54) take out its ${}^{i}_{j}$ components

$$\begin{cases} D_{\mu}\phi_{j}^{i} = \partial_{\mu}\phi_{j}^{i} - \mathrm{i}gA_{\mu k}^{i}\phi_{j}^{k} + \mathrm{i}gA_{\mu j}^{l}\phi_{l}^{i} \\ D^{\mu}\phi_{i}^{j} = \partial^{\mu}\phi_{i}^{j} + \mathrm{i}gA_{k}^{\mu j}\phi_{i}^{k} - \mathrm{i}gA_{i}^{\mu l}\phi_{l}^{j} \end{cases} \tag{11.1.55}$$

$$\left\langle D^{\mu}\phi_{i}^{j}D_{\mu}\phi_{j}^{i}\right\rangle = -g^{2}\left(\omega_{i}-\omega_{j}\right)^{2}A_{i}^{\mu j}A_{\mu j}^{i} \tag{11.1.56}$$

where Eq. (11.1.27) leads to

$$\omega_1 = \omega_2 = \omega_3 = -\frac{2}{3}a, \quad \omega_4 = \omega_5 = a \tag{11.1.57}$$

Obviously, when $i,j = 1,2,3$ or $i,j = 4,5$, the corresponding gauge bosons do not obtain masses at the first step of the symmetry breaking, i.e. gluons $(i,j = 1,2,3)$ and photon, $\mathrm{W}^{\pm}$ and $\mathrm{Z}^{0}(i,j = 4,5)$ still keep massless. But from Eq. (11.1.25) we see that for $i = 1$, $j = 4$ we obtain the X boson mass, while for $i = 1$, $j = 5$ we get the Y boson mass, i.e.

$$m_{\mathrm{X}}^{2} = m_{\mathrm{Y}}^{2} = \frac{1}{2}\cdot\frac{25}{9}g^{2}a^{2} \tag{11.1.58}$$

Because of Eq. (11.1.25), only $\hat{A}_{\mu}$ is the gauge boson with the right normalization, while A_{μ} is not. Using Eq. (11.1.22), $A_{\mu} = \frac{1}{\sqrt{2}}\hat{A}_{\mu}$, the mass term should use $\hat{A}_{j}^{\mu_i}\hat{A}_{\mu_i}^{j}$, so in Eq. (11.1.58) it appears the factor $\frac{1}{2}$.

Now let us look at the second step of symmetry breaking. The second step is $\mathrm{SU(3)_C}\times\mathrm{SU(2)}\times\mathrm{U(1)} \xrightarrow{10^2\ \mathrm{GeV}} \mathrm{SU(3)_C}\times\mathrm{U(1)_{em}}$, the corresponding gauge fields are as follows:

$$G_\mu = \frac{1}{2}T_\alpha G_\mu^\alpha = \begin{pmatrix} G_\mu & 0 \\ 0 & \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} \end{pmatrix} \tag{11.1.59}$$

$$W_\mu = \frac{1}{2}\vec{T}\cdot\vec{W}_\mu = \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{2}\vec{\sigma}\cdot\vec{W} \end{pmatrix} \tag{11.1.60}$$

$$E_\mu = \frac{1}{2}YB_\mu = \frac{1}{2}\sqrt{\frac{3}{5}}B_\mu \begin{pmatrix} -2/3 & & & & \\ & -2/3 & & & \\ & & -2/3 & & \\ & & & 1 & \\ & & & & 1 \end{pmatrix} \tag{11.1.61}$$

$$\begin{cases} D_\mu\varphi^i = \partial_\mu\varphi^i - \mathrm{i}g_3 G^i_{\mu k}\varphi^k - \mathrm{i}g_2 W^i_{\mu k}\varphi^k - \mathrm{i}g_1 E^i_{\mu k}\varphi^k \\ (D^\mu\varphi^i)^+ = \partial^\mu\varphi_i + \mathrm{i}g_3 G_i^{\mu l}\varphi_l + \mathrm{i}g_2 W_i^{\mu l}\varphi_l + \mathrm{i}g_1 E_i^{\mu l}\varphi_l \end{cases} \tag{11.1.62}$$

Using $\langle\varphi^i\rangle = \upsilon\delta_5^i$, $\mathrm{W}^\pm = \frac{W_1 \mp \mathrm{i}W_2}{\sqrt{2}}$, then

$$\begin{aligned} \langle(D^\mu\varphi^i)^+(D_\mu\varphi^i)\rangle = \upsilon^2 \Bigg\{ & \frac{1}{2}g_2^2 W_\mu^- W^{+\mu} + \frac{1}{4}g_2^2 W_\mu^3 W^{3\mu} \\ & + \frac{1}{4}g_1^2\frac{3}{5}B^\mu B_\mu + \frac{1}{4}g_1 g_2\sqrt{\frac{3}{5}}(-2)B^\mu W_\mu^3 \Bigg\} \end{aligned} \tag{11.1.63}$$

Thus, the gluons still keep massless, i.e.

$$m_{\mathrm{G}}^2 = 0 \tag{11.1.64}$$

But

$$m_{\mathrm{W}^\pm}^2 = \frac{1}{2}g_2^2\upsilon^2 \tag{11.1.65}$$

W^3 and B have mixing, their mass mixing matrix is

$$\frac{\upsilon^2}{4}\begin{pmatrix} g_2^2 & -\sqrt{\frac{3}{5}}g_1 g_2 \\ -\sqrt{\frac{3}{5}}g_1 g_2 & \frac{3}{5}g_1^2 \end{pmatrix} \tag{11.1.66}$$

Diagonalizing Eq. (11.1.66) we get the corresponding eigenvalues and eigenvectors

$$\begin{cases} \lambda_1 = 0, & \text{eigenvector}\, A_\mu = B_\mu \cos\theta_{\mathrm{W}} + W_\mu^3 \sin\theta_{\mathrm{W}} \\ \lambda_2 = \left(\dfrac{3}{5}g_1^2 + g_2^2\right)\dfrac{\upsilon^2}{4}, & \text{eigenvector}\, Z_\mu = -B_\mu \sin\theta_{\mathrm{W}} + W_\mu^3 \cos\theta_{\mathrm{W}} \end{cases} \tag{11.1.67}$$

where

$$\tan\theta_{\mathrm{W}} = \sqrt{\frac{3}{5}}g_1/g_2 \tag{11.1.68}$$

so photon field A_μ is still massless, but Z_μ gets mass:

$$\begin{cases} m_\gamma^2 = 0 \\ m_{\mathrm{Z}}^2 = \left(\dfrac{3}{5}g_1^2 + g_2^2\right)\dfrac{\upsilon^2}{2} \end{cases} \tag{11.1.69}$$

$$\frac{m_{\mathrm{W}^\pm}^2}{m_{\mathrm{Z}}^2} = \frac{g_2^2}{g_2^2 + \frac{3}{5}g_1^2} = \frac{1}{1+\tan^2\theta_{\mathrm{W}}} = \cos^2\theta_{\mathrm{W}} \tag{11.1.70}$$

These are just the relations in SU(2) × U(1) electroweak unified model. From Eq. (11.1.70), it is easy to derive

$$\sin^2\theta_{\mathrm{W}} = \frac{g_1^2}{g_1^2 + \frac{5}{3}g_2^2} \tag{11.1.71}$$

At the grand unification energy scale M, $g_1^2(M) = g_2^2(M) = g_3^2(M) = g^2$,

$$\sin^2\theta_{\mathrm{W}}\big|_M = \sin^2\theta_{\mathrm{W}}(M) = 3/8 \tag{11.1.72}$$

To get the value of $\sin^2\theta_{\mathrm{W}}$ at low energy $M_0 = M_{\mathrm{Z}}$, we can use renormalization group equation

$$\begin{cases} \dfrac{\mathrm{d}g_i(\mu)}{\mathrm{d}t} = \beta_i(g_j) = b_i g_i^3 + O(g_i^5) \\ t = \ln\mu, \quad i,j = 1,2,3 \end{cases} \tag{11.1.73}$$

Neglecting terms of the order of g_i^5,

$$\frac{\mathrm{d}g_i(\mu)}{\mathrm{d}t} = b_i g_i^3 \tag{11.1.74}$$

Integrating Eq. (11.1.74)

$$\frac{1}{g_i^2(M_0)} = \frac{1}{g_i^2(M)} + 2b_i \ln(M/M_0) \tag{11.1.75}$$

The values of the $b_i(i,j=1,2,3)$ are

$$\left.\begin{aligned}&\text{for}\quad \mathrm{U}(1),\quad b_1=\frac{1}{48\pi^2}2n_\mathrm{f}\\&\text{for}\quad \mathrm{SU}(2),\quad b_2=\frac{1}{48\pi^2}(-11\times 2+2n_\mathrm{f})\\&\text{for}\quad \mathrm{SU}(3),\quad b_3=\frac{1}{48\pi^2}(-11\times 3+2n_\mathrm{f})\end{aligned}\right\}\tag{11.1.76}$$

where n_f is the number of fermion flavors with mass $\leq M$.

Now we look at the relations of g_2, g_1 with g, g' in $\mathrm{SU}(2)\times\mathrm{U}(1)$ electroweak unified theory.

At the second step breaking of SU(5), $\mathrm{SU}(3)_\mathrm{C}\times\mathrm{SU}(2)\times\mathrm{U}(1)$ is broken down to $\mathrm{SU}(3)_\mathrm{C}\times\mathrm{U}(1)_\mathrm{em}$, but $\mathrm{SU}(3)_\mathrm{C}$ symmetry is kept. While $\mathrm{SU}(2)\times\mathrm{U}(1)$ is broken, its covariant derivative becomes to

$$D_\mu=\partial_\mu-\mathrm{i}g_2\frac{\vec{T}}{2}\cdot\underset{\mu}{\vec{W}}-\mathrm{i}g_1\frac{Y}{2}B_\mu\tag{11.1.77}$$

But in electroweak unified theory, the covariant derivative is

$$D_\mu=\partial_\mu-\mathrm{i}g\,\vec{I}\cdot\vec{W}-\mathrm{i}g'(Q-I_3)B_\mu\tag{11.1.78}$$

Comparing with Eq. (11.1.77), we find that

$$\frac{T}{2}=\begin{pmatrix}0&0\\0&\frac{1}{2}\vec{\sigma}\end{pmatrix}=\vec{I}\tag{11.1.79}$$

Thus

$$g_2=g\tag{11.1.80}$$

While

$$\begin{cases}I_3=\mathrm{diag}\,(0,0,0,1/2,-1/2)\\Q=\mathrm{diag}(-1/3-1/3,-1/3,1,0)\end{cases}\tag{11.1.81}$$

$$Q-I_3=\mathrm{diag}(-1/3,-1/3,-1/3,1/2,1/2)=\sqrt{\frac{5}{3}}\frac{Y}{2}\tag{11.1.82}$$

Comparing Eqs. (11.1.77), (11.1.78) and (11.1.82) we get

$$g_1=\sqrt{\frac{5}{3}}g'\tag{11.1.83}$$

So comparing the electroweak unified theory with the second step breaking of SU(5) GUT, we obtain

$$e(M_0) = g_2(M_0)\sin\theta_{\mathrm{W}}(M_0) = g(M_0)\sin\theta_{\mathrm{W}}(M_0) = g'(M_0)\cos\theta_{\mathrm{W}}(M_0) \tag{11.1.84}$$

This leads to

$$\frac{g_2^2(M_0)}{4\pi} = \frac{\alpha(M_0)}{\sin^2\theta_{\mathrm{W}}(M_0)} \tag{11.1.85}$$

where

$$\alpha(M_0) = \frac{e^2(M_0)}{4\pi} \tag{11.1.86}$$

Substituting Eq. (11.1.83) into Eq. (11.1.85) and using Eq. (11.1.84), we get

$$\frac{g_1^2(M_0)}{4\pi} = \frac{5}{3}\frac{\alpha(M_0)}{\cos^2\theta_{\mathrm{W}}(M_0)} \tag{11.1.87}$$

Define

$$\alpha_s(M_0) = \frac{g_3^2(M_0)}{4\pi} \tag{11.1.88}$$

Substituting Eq. (11.1.76) into Eq. (11.1.75) and eliminating n_{f} and $\ln M/M_0$, we can get

$$\sin^2\theta_{\mathrm{W}}(M_0) = \frac{1}{6} + \frac{5}{9}\frac{\alpha(M_0)}{\alpha_s(M_0)} \tag{11.1.89}$$

If eliminating n_{f} and $\sin^2\theta_{\mathrm{W}}$, then we have

$$\frac{\alpha(M_0)}{\alpha_s(M_0)} = \frac{3}{8} - \frac{33}{8\pi}\alpha(M_0)\ln(M/M_0) \tag{11.1.90}$$

From the above two equations, we can derive

$$\sin^2\theta_{\mathrm{W}}(M_0) = \frac{3}{8} - \frac{55}{24\pi}\alpha(M_0)\ln(M/M_0) \tag{11.1.91}$$

Utilizing $M \sim 10^{15}\,\mathrm{GeV}$, $M_0 \sim 100\,\mathrm{GeV}$, $\alpha(M_0) \sim 1/128.5$, Eq. (11.1.91) leads to

$$\sin^2\theta_{\mathrm{W}}(100\,\mathrm{GeV}) \sim 0.2 \tag{11.1.92}$$

This is compatible with the experimental value [5]

$$\sin^2\theta_{\mathrm{W}}(M_Z) = 0.23116 \pm 0.00018 \tag{11.1.93}$$

11.1.1.6. *Proton decay*

Because fermions fill up $\underline{10}_{\mathrm{L}}+\underline{5}_{\mathrm{R}}$ representation, we see that quarks and leptons are in the same representation so can transform into each other. This transformation will induce proton decay. Of course, proton decay lifetime must be long enough, or else all things in the Universe would disappear too fast, and this is not consistent with experimental facts.

In the SU(5) GUT model, proton decay is induced by the super-heavy gauge bosons $\mathrm{X}(|Q|=4/3)$, $\mathrm{Y}(|Q|=1/3)$ and super-heavy Higgs particles. It can be proved that [3] the decays induced by super-heavy Higgs are slower than that by super-heavy X, Y, so can be neglected. The proton decay Hamiltonian induced by X, Y is

$$\begin{aligned} L_i = {} & \frac{1}{\sqrt{2}} g X^-_{i\mu}(-\bar{d}_{i\mathrm{R}}\gamma^\mu e^+_{\mathrm{R}} + \varepsilon_{ijk}\bar{u}^c_{k\mathrm{L}}\gamma^\mu u_{j\mathrm{L}} + \bar{d}_{i\mathrm{L}}\gamma^\mu e^+_{\mathrm{L}}) \\ & + \frac{1}{\sqrt{2}} g Y^-_{i\mu}(\bar{d}_{i\mathrm{R}}\gamma^\mu \nu^c_{\mathrm{eR}} + \varepsilon_{ijk}\bar{u}^c_{k\mathrm{L}}\gamma^\mu d_{j\mathrm{L}} + \bar{u}_{i\mathrm{L}}\gamma^\mu e^+_{\mathrm{L}}) + \text{h.c.} \end{aligned} \tag{11.1.94}$$

The corresponding Feynman diagrams are sketched in Fig. 11.2.

The careful calculation in the minimal SU(5) GUT model gives [6] the proton lifetime

$$\begin{cases} \tau_{\mathrm{p}} \approx 2\times 10^{29\pm 1.7}\,\mathrm{yr} \\ \tau_{\mathrm{p}\to \mathrm{e}^+\pi^0} = 1/\varGamma(p\to \mathrm{e}^+\pi^0) = 4.5\times 10^{29\pm 1.7}\,\mathrm{yr} \end{cases} \tag{11.1.95}$$

where $\mathrm{p}\to \mathrm{e}^+\pi^0$ is the dominant decay channel for the minimal SU(5) GUT model.

The latest measured values by super-Kamiokande experiments are [7] as follows:

$$\begin{cases} \tau_{\mathrm{p}\to \mathrm{e}^+\pi^0} = 1/\varGamma(p\to \mathrm{e}^+\pi^0) > 1.0\times 10^{34}\,\mathrm{yr} \\ \tau_{\mathrm{p}\to \mathrm{K}^+\bar{\nu}} = 1/\varGamma(p\to \mathrm{K}^+\bar{\nu}) > 3.3\times 10^{33}\,\mathrm{yr} \quad \text{at } 90\%\mathrm{CL} \\ \tau_{\mathrm{n}\to \mathrm{e}^+\pi^-} = 1/\varGamma(n\to \mathrm{e}^+\pi^-) > 2\times 10^{33}\,\mathrm{yr} \end{cases} \tag{11.1.96}$$

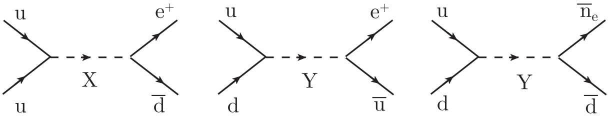


Fig. 11.2. Feynman diagrams for proton decay.

Comparing Eq. (11.1.95) with Eq. (11.1.96), we see that the minimal SU(5) GUT model is ruled out. Certainly, non-minimal SU(5) GUT models (for example, introduce three 5-dimensional Higgs $\phi^i_\alpha (\alpha = 1, 2, 3)$ and one 45-dimensional Higgs ϕ^i_{jk}) are still not ruled out.

11.1.2. *SO(2N) grand unified models*

The first model discussed in detail in the literature is SO(10). In the minimal SO(10) GUT [8] model, the fermions are filled in the 16-dimensional spinor representation, its decomposition with respect to SU(5) is

$$\underline{16} = [10 + \bar{5} + 1] \tag{11.1.97}$$

So, the right-handed component of the neutrino missing in SU(5) GUT model now can be filled in the singlet in the above decomposition of Eq. (11.1.97); this makes neutrino to obtain Dirac mass naturally.

The other advantage of the SO (10)GUT model is that the calculated proton lifetime is not contradictory with experiments. Its disadvantage is the same as SU(5) GUT, that is, the three fermion generations can only fill in the same 16-dimensional spinor representation repeatedly.

The spontaneous symmetry breaking mechanism is more complicated than SU(5) GUT, it has different ways to be broken into the same $\mathrm{SU}(3)_\mathrm{C} \times \mathrm{U}(1)_\mathrm{em}$. Here we shall not discuss it in detail, the readers who are interested in it can refer the relative references.

We need to point out that SO(10) model is not in conflict with proton decay experiments, so it is not ruled out. In the literature there are discussions to the SO(14) and SO(18) GUT models, it will be mentioned again latter.

11.1.3. *Flavor unification*

We have seen in above subsections that for SU(5) and SO(10) GUT models the three fermion generations can only fill in one representation repeatedly. These repeated filling is unnatural, because the three generations of quarks and leptons are not unified in one group representation. This is so-called generation problem.

In order to solve this problem, Georgi proposed three principles for constructing $\mathrm{SU}(N)\,(N > 5)$ GUT models:

(1) left-handed fermion representation must be real with respect to $\mathrm{SU}(3)_\mathrm{C}$ subgroup, so as to guarantee $\mathrm{SU}(3)_\mathrm{C}$ symmetry not broken;

(2) left-handed fermion representation should be complex with respect to the subgroup $\mathrm{SU}(3)_\mathrm{C} \times \mathrm{SU}(2) \times \mathrm{U}(1)$;
(3) in left-handed fermion representations any irreducible representation can appear only once.

According to the above three principles, only $\mathrm{SU}(N)$ for $N \geq 11$ can accommodate three fermion generations at the same time.

For the above three principles (1) and (3) cannot be abandoned, but (2) can be revised appropriately so as to use lower rank group for accommodating three fermion generations.

In fact, the first step breaking of GUT gauge group into $G_1 = \mathrm{SU}(3)_\mathrm{C} \times \mathrm{SU}(2) \times \mathrm{U}(1)$ can be changed. For example, the first step breaking could be into $G_1' = \mathrm{SU}(3)_\mathrm{C} \times \mathrm{SU}(2) \times \mathrm{U}(1) \times \mathrm{U}(1)'$ or $G_1' = \mathrm{SU}(3)_\mathrm{C} \times \mathrm{SU}(2) \times \mathrm{U}(1) \times G_D$, G_D is a discrete symmetry group, then the lower rank group can also accommodate three fermion generations. That is, light fermion representation can be real with respect to $G_1 = \mathrm{SU}(3)_\mathrm{C} \times \mathrm{SU}(2) \times \mathrm{U}(1)$, but it is complex with respect to

$$G_1' = \mathrm{SU}(3)_\mathrm{C} \times \mathrm{SU}(2) \times \mathrm{U}(1) \times \begin{cases} \mathrm{U}(1)' \\ G_D \end{cases}, \tag{11.1.98}$$

In this way, we still can prevent light fermions from getting super-heavy mass at the first step symmetry breaking. According to this idea SU(7) GUT can accommodate two generations of light fermions [9], SU(8) GUT can accommodate three generations of light fermions with V–A weak coupling interactions, and two generations of light fermions with V+A coupling weak interactions and half generation light fermions with V–A and V + A mixed coupling [10].

In the same way, we can construct SO(14) [11], SO(18) [12] GUT models which can accommodate more generations of light fermions.

Of course, there are still more other GUT models, we shall not discuss them here. The readers can refer the relative references.

11.2. Supersymmetric Models

Another important direction beyond standard model is supersymmetry. Supersymmetry is a symmetry between fermions and bosons. Firstly, introducing supersymmetry is from esthetics point of view, i.e. every fermion should correspond to a boson and vice versa. In addition, from technology point of view introducing supersymmetry (SUSY) can make the gauge mass hierarchy not be affected from radiative corrections. In fact, in the standard model one loop radiative correction of the Higgs mass is quadratically divergent, i.e. the Higgs mass correction is proportional to the cutoff energy scale squired. That is certainly disaster, because in the standard model the masses of fermions and bosons (gauge bosons) are all connected with the vacuum expectation value of the Higgs particle, so are not stable at one loop radiative corrections. But after introducing SUSY, the quadratic divergence of the scalar one loop radiative corrections is canceled by its corresponding superpartner (fermion) one loop radiative corrections, so the mass gauge hierarchy is kept. This kind of divergence cancellation between fermion loop radiative corrections and its corresponding boson loop radiative corrections is the typical feature of SUSY models, and is also one of the reasons why we introduce supersymmetry.

There is a third reason for introducing SUSY. Usually in GUT models (for example, SU(5) GUT) the coupling constants of the strong, electromagnetic and weak interactions should join together in one point at the GUT energy scale ($\sim 10^{15}$ GeV) so as to unify into one interaction. But if we look at this microscopically, the three couplings do not really join at one point. But after introducing SUSY, the three couplings can really join together at one point. (see Fig. 11.3) [13].

Certainly, this reason is not compulsory, if you think that the three interactions do not need to unify, so do not need to join at one point.

As we discussed above, SUSY transforms fermions into bosons and vice versa, so the SUSY generators should be fermion-type operators, and carry

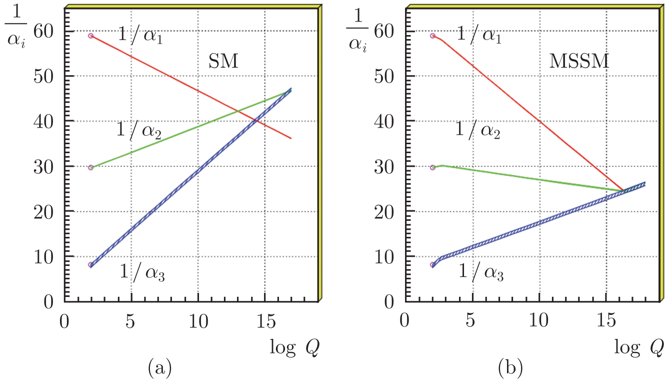


Fig. 11.3. Evolution of coupling constants along with energy scale: (a) GUT without SUSY; (b) GUT with SUSY; experimental values of coupling constants are from LEP measurements in 1991; the difference of the two figures is that the figure on the right-hand side includes the contribution of the superpartner particles.

spin quantum number 1/2 and obey anti-commutation relations. We use Weyl spinor operator $Q_\alpha (\alpha = 1, 2)$ to express them. Obviously,

$$Q\left|\text{Boson}\right\rangle = \left|\text{Fermion}\right\rangle, \quad Q\left|\text{Fermion}\right\rangle = \left|\text{Boson}\right\rangle \tag{11.2.1}$$

Q_α transform under Lorentz transformation according to $(1/2, 0)$ representation. Put

$$\bar{Q} \equiv Q^{+} \tag{11.2.2}$$

then $\left\{Q_\alpha, \bar{Q}_{\dot{\alpha}}\right\}$ transform according to Lorentz group representation $(1/2, 1/2)$. From the Coleman–Mandula theorem [14], SUSY transformations allow energy–momentum P^μ to conserve, so we have

$$\begin{cases} [Q_\alpha^A, P^\mu] = [\bar{Q}_{\dot{\beta}B}, P^\mu] = 0 \\ \{Q_\alpha^A, \bar{Q}_{\dot{\beta}B}\} = 2\sigma^\mu_{\alpha\dot{\beta}} P_\mu \delta^A_B \\ \{Q_\alpha^A, Q_\beta^B\} = \{\bar{Q}_{\dot{\alpha}A}, \bar{Q}_{\dot{\beta}B}\} = 0 \end{cases} \tag{11.2.3}$$

where $\alpha, \beta, \ldots, \dot{\alpha}, \dot{\beta}, \ldots$ take 1 or 2, $\mu = 0, 1, 2, 3$. $\sigma_\mu = (1, \overrightarrow{\sigma})$, i.e. $\sigma_0 = 1$, $\sigma_i\,(i = 1, 2, 3)$ are Pauli matrices. The space–time matrix $g_{\mu\nu} = g^{\mu\nu} = \text{diag}(1, -1, -1, -1)$. $A, B = 1, 2, \ldots, N, \ldots$, the corresponding SUSY are called $N = 1$, $N = 2, \ldots, N = N, \ldots$ SUSY.

In this book, we only discuss $N = 1$ SUSY, so only the special case of Eq. (11.2.3) is involved, i.e. we only need to consider the following case:

$$\begin{cases} [Q_\alpha, P^\mu] = [\bar{Q}_{\dot{\beta}}, P^\mu] = 0 \\ \{Q_\alpha, \bar{Q}_{\dot{\beta}}\} = 2\sigma^\mu_{\alpha\dot{\beta}} P_\mu \\ \{Q_\alpha, Q_\beta\} = \{\bar{Q}_{\dot{\alpha}}, \bar{Q}_{\dot{\beta}}\} = 0 \end{cases} \tag{11.2.4}$$

The common symbol rule in SUSY can be found in [15].

From the second equation of Eq. (11.2.4), it is easy to see that

$$H \equiv P^0 = \frac{1}{4}\left(\bar{Q}_1 Q_1 + Q_1 \bar{Q}_1 + \bar{Q}_2 Q_2 + Q_2 \bar{Q}_2\right) \tag{11.2.5}$$

If the vacuum $|0\rangle$ is SUSY invariant, i.e.

$$Q_\alpha\left|0\right\rangle = 0, \quad \left\langle 0\right|\bar{Q}_\alpha = 0 \tag{11.2.6}$$

then we can have

$$\left\langle 0\right| H \left|0\right\rangle = E_{\text{vac}} = 0 \tag{11.2.7}$$

i.e. the vacuum energy E_{vac} is always the absolute energy minimum.

But if SUSY is spontaneously broken, i.e. the Lagrangian is SUSY invariant, but the vacuum is not, that will make

$$Q_\alpha \left|0\right\rangle \neq 0, \quad \left\langle 0\right|\bar{Q}_\alpha \neq 0 \tag{11.2.8}$$

This means that the vacuum energy $E_{\text{vac}} > 0$. This can be seen from Eq. (11.2.5)

$$E_{\text{vac}} = \left\langle 0\right| H \left|0\right\rangle > 0 \tag{11.2.9}$$

Thus, if SUSY is spontaneously broken, there must be $E_{\text{vac}} > 0$.

Because up to now we still not find super partners in experiments, so SUSY must be broken at low energy. But SUSY breaking is not easily realized. So in commonly used SUSY models in particle physics, for instance, the minimal supersymmetric standard model (MSSM), we generally introduce soft supersymmetry breaking artificially to break SUSY. The most attractive spontaneously SUSY breaking models are very complicated and not easy to construct a realistic model.

Owing to space limitation, in this book we only introduce the basic concepts of SUSY, many formulas will not give derivations. The readers who want to work in SUSY field should refer the original references. You can start with the following reviews [16].

11.2.1. *Superspace and superfields*

Based on the space–time coordinates t, x, y, z and adding four anti-commutating Grassmann parameters θ^α, $\bar{\theta}^{\dot{\beta}} (\alpha, \dot{\beta} = 1, 2)$ a superspace manifold is constructed. Discussing SUSY in superspace is more convenient and simple, where

$$\left\{\theta^\alpha, \theta^\beta\right\} = \left\{\bar{\theta}^{\dot{\alpha}}, \bar{\theta}^{\dot{\beta}}\right\} = \left\{\theta^\alpha, \bar{\theta}^{\dot{\beta}}\right\} = 0 \tag{11.2.10}$$

In superspace, SUSY algebra can be written as

$$\begin{cases} \left[\theta Q, \bar{Q}\bar{\theta}\right] = 2\theta\sigma_\mu\bar{\theta}P^\mu \\ \left[\theta Q, \theta Q\right] = \left[\bar{Q}\bar{\theta}, \bar{Q}\bar{\theta}\right] = 0 \end{cases} \tag{11.2.11}$$

where σ_μ is already defined before, i.e.

$$\begin{cases} \sigma_\mu = \left(1, \vec{\sigma}\right), \quad \text{i.e.} \quad \sigma_0 = 1, \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} \\ \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \text{ are Pauli matrices} \end{cases} \tag{11.2.12}$$

$$\theta Q = \theta^\alpha Q_\alpha = \theta^\alpha Q^\beta \varepsilon_{\alpha\beta}, \quad \varepsilon_{12} = 1, \quad \varepsilon_{21} = -1 \tag{11.2.13}$$

A finite SUSY transformation is defined as

$$S(x, \theta, \bar{\theta}) = \mathrm{e}^{\mathrm{i}(\theta Q + \bar{Q}\bar{\theta} + x_\mu P^\mu)} \tag{11.2.14}$$

Two successive SUSY transformations are given according to Hausdorff formula:

$$S(x, \theta, \bar{\theta}) S(y, \alpha, \bar{\alpha}) = S(x + y - \mathrm{i}\alpha\sigma_\mu\bar{\theta} + \mathrm{i}\theta\sigma_\mu\bar{\alpha},\ \theta + \alpha,\ \bar{\theta} + \bar{\alpha}) \tag{11.2.15}$$

Any superfield $\phi(x, \theta, \bar{\theta})$ is defined as a function of $x, \theta, \bar{\theta}$ and satisfies the following SUSY transformation relations:

$$S\left(y, \alpha, \bar{\alpha}\right) \phi\left(x, \theta, \bar{\theta}\right) = \phi\left(x + y - \mathrm{i}\alpha\sigma_\mu\bar{\theta} + \mathrm{i}\theta\sigma_\mu\bar{\alpha},\ \theta + \alpha,\ \bar{\theta} + \bar{\alpha}\right) \tag{11.2.16}$$

We shall discuss the superfield properties later.

11.2.2. *Global supersymmetry and local supersymmetry (Supergravity)*

If in a super–transformation $S(x, \theta, \bar{\theta})$ the superspace parameters θ, $\bar{\theta}$ are independent of the space–time coordinates x_μ, we call $S(x, \theta, \bar{\theta})$ as a global SUSY transformation. A theory invariant under global SUSY transformation is called global SUSY theory. But if the above SUSY transformation parameters θ, $\bar{\theta}$ depend on the space–time coordinates, i.e. $\theta(x)$, $\bar{\theta}(x)$, then this SUSY transformation $S\left(x, \theta(x), \bar{\theta}(x)\right)$ is called local SUSY transformation. A theory invariant under local SUSY transformations is called local SUSY theory. Because θ, $\bar{\theta}$ depend on the space–time coordinates x_μ, the gravity must appear in this theory. So local SUSY theory is also called supergravity theory. Actually, in references today global supersymmetric theory is called simply SUSY theory and neglect the word “global”. While local supersymmetric theory is called simply supergravity theory. In particle physics, frequently-used theory is global SUSY theory.

11.2.3. *Global supersymmetric theory and minimal supersymmetric standard model*

11.2.3.1. *Scalar (chiral) superfield and vector superfield*

For any superfield $\phi\left(x,\theta,\bar{\theta}\right)$, Eq. (11.2.16) already gives the transformation formula under supersymmetric transformation. For an infinitesimal supertransformation $S(0,\varepsilon,\bar{\varepsilon})$, Eq. (11.2.14) leads to

$$S\left(0,\varepsilon,\bar{\varepsilon}\right)=\mathrm{e}^{\mathrm{i}\left(\varepsilon Q+\bar{Q}\bar{\varepsilon}\right)} \tag{11.2.17}$$

Equation (11.2.16) becomes to

$$S(0,\varepsilon,\bar{\varepsilon})\phi(x,\theta,\bar{\theta})=\phi(x-\mathrm{i}\varepsilon\sigma_\mu\theta+\mathrm{i}\theta\sigma_\mu\bar{\varepsilon},\theta+\varepsilon,\bar{\theta}+\bar{\varepsilon}) \tag{11.2.18}$$

$$\begin{aligned}\delta\varphi(x,\theta,\bar{\theta})&=\varphi(x-\mathrm{i}\varepsilon\sigma_\mu\bar{\theta}+\mathrm{i}\theta\sigma_\mu\bar{\varepsilon},\theta+\varepsilon,\bar{\theta}+\bar{\varepsilon})-\varphi(x,\theta,\bar{\theta})\\&=\left[\varepsilon\frac{\partial}{\partial\theta}+\bar{\varepsilon}\frac{\partial}{\partial\bar{\theta}}-\mathrm{i}(\varepsilon\sigma_\mu\bar{\theta}-\theta\sigma_\mu\bar{\varepsilon})\partial^\mu\right]\varphi(x,\theta,\bar{\theta})\end{aligned} \tag{11.2.19}$$

From the above formulas, we can derive

$$\begin{cases}P_\mu=-\mathrm{i}\dfrac{\partial}{\partial x^\mu}=-\mathrm{i}\partial_\mu\\ Q_\alpha=\dfrac{\partial}{\partial\theta^\alpha}-\mathrm{i}\sigma^\mu_{\alpha\dot{\beta}}\bar{\theta}^{\dot{\beta}}\partial_\mu\\ \bar{Q}_{\dot{\alpha}}=-\dfrac{\partial}{\partial\bar{\theta}^{\dot{\alpha}}}+\mathrm{i}\theta^\beta\sigma^\mu_{\beta\dot{\alpha}}\partial_\mu\end{cases} \tag{11.2.20}$$

In superspace we can define the covariant derivatives D_α, $\bar{D}_{\dot{\alpha}}$ and demand them to anti-commute with supertransformation, i.e.

$$D_\alpha(\delta\phi)=-\delta(D_\alpha\phi) \tag{11.2.21}$$

Then we have

$$\begin{cases}D_\alpha=\left(\dfrac{\partial}{\partial\theta^\alpha}+\mathrm{i}\sigma^\mu_{\alpha\dot{\beta}}\bar{\theta}^{\dot{\beta}}\dfrac{\partial}{\partial x^\mu}\right)\\ \bar{D}_{\dot{\alpha}}=\left(-\dfrac{\partial}{\partial\bar{\theta}^{\dot{\alpha}}}-\mathrm{i}\theta^\beta\sigma^\mu_{\beta\dot{\alpha}}\dfrac{\partial}{\partial x^\mu}\right)\end{cases} \tag{11.2.22}$$

We can define the generators and the covariant derivatives under left-handed representation and right-handed representation. For left-handed representation, they are

$$\begin{cases} Q_{\mathrm{L}} = \dfrac{\partial}{\partial\theta}, \quad \bar{Q}_{\mathrm{L}} = -\dfrac{\partial}{\partial\bar{\theta}} + 2\mathrm{i}\theta\sigma_\mu\partial^\mu \\ D_{\mathrm{L}} = \dfrac{\partial}{\partial\theta} + 2\mathrm{i}\sigma_\mu\bar{\theta}\partial^\mu, \quad \bar{D}_{\mathrm{L}} = -\dfrac{\partial}{\partial\bar{\theta}} \end{cases} \tag{11.2.23}$$

for right-handed representation they are

$$\begin{cases} Q_{\mathrm{R}} = \dfrac{\partial}{\partial\theta} - 2\mathrm{i}\sigma^\mu\bar{\theta}\partial_\mu, \quad \bar{Q}_{\mathrm{R}} = -\dfrac{\partial}{\partial\bar{\theta}} \\ D_{\mathrm{R}} = \dfrac{\partial}{\partial\theta}, \quad \bar{D}_{\mathrm{R}} = -\dfrac{\partial}{\partial\bar{\theta}} - 2\mathrm{i}\theta\sigma_\mu\partial^\mu \end{cases} \tag{11.2.24}$$

We call a superfield ϕ with $\bar{D}\phi = 0$ or $D\phi = 0$ as scalar superfield or chiral superfield. But sometimes we call a superfield ϕ satisfying $\bar{D}\phi = 0$ as left-handed chiral superfield, while a superfield ϕ satisfying $D\phi = 0$ as right-handed chiral superfield.

Generally speaking, scalar superfield or left-handed (right-handed) chiral superfield are all functions of $(x, \theta, \bar{\theta})$. But at some special case these functions can be simplified. For example, at left-handed representation, using Eq. (11.2.23),

$$\bar{D}_{\mathrm{L}}\phi_{\mathrm{L}} = 0 \tag{11.2.25}$$

Then ϕ_{L} is only function of (x, θ), this can be seen from Eq. (11.2.25). Thus

$$\phi_{\mathrm{L}}(x, \theta) = \varphi(x) + \theta^\alpha\psi_\alpha(x) + \theta^\alpha\theta^\beta\varepsilon_{\alpha\beta}F(x) \tag{11.2.26}$$

Note that θ is a Grassmann number, so terms with $\theta^n (n \geq 3)$ are all zero, and $\varphi(x)$, $F(x)$ are all complex scalar fields. If we take mass M as the dimension scale, the dimension of θ is $M^{-1/2}$, the dimension of $F(x)$ is M^2, that of $\varphi(x)$ is M.

Under infinitesimal SUSY transformation

$$\delta\phi_{\mathrm{L}}(x, \theta) = \delta\varphi(x) + \theta\delta\psi(x) + \theta\theta\delta F(x) \tag{11.2.27}$$

where

$$\begin{cases} \delta\varphi = \sqrt{2}\varepsilon\psi \\ \delta\psi = \sqrt{2}\varepsilon F + \mathrm{i}\sqrt{2}\sigma_\mu\bar{\varepsilon}\partial^\mu\varphi \\ \delta F = -\mathrm{i}\sqrt{2}\partial^\mu\psi\sigma_\mu\bar{\varepsilon} \end{cases} \tag{11.2.28}$$

From Eq. (11.2.28) we see that the change of the F term under the SUSY transformation is a total derivative. This characteristic can be used to construct Lagrangian (see the following section). We can see also that SUSY transformation indeed changes fermion into boson, and boson into fermion.

In Eq. (11.2.26) the chiral superfield includes only spin zero scalar fields φ, F, and spin 1/2 spinor field ψ. We also need to introduce spin-1 vector superfield $V(x,\theta,\bar{\theta})$. If we put self-conjugation condition $V^{+}(x,\theta,\bar{\theta}) = V(x,\theta,\bar{\theta})$ on general superfield, then we can get vector superfield. Its general form is

$$\begin{aligned} V(x,\theta,\bar{\theta}) = & \left(1+\frac{1}{4}\theta\theta\bar{\theta}\bar{\theta}\Box\right)C(x) \\ & +(\mathrm{i}\theta+\frac{1}{2}\theta\theta\sigma^{\mu}\bar{\theta}\partial_{\mu})\chi(x)+\frac{1}{2}\mathrm{i}\theta\theta(M(x)+\mathrm{i}N(x)) \\ & +(-\mathrm{i}\bar{\theta}+\frac{1}{2}\bar{\theta}\bar{\theta}\theta\sigma_{\mu}\partial^{\mu})\bar{\chi}(x)-\frac{1}{2}\mathrm{i}\bar{\theta}\bar{\theta}(M(x)-\mathrm{i}N(x)) \\ & -\theta\partial_{\mu}\bar{\theta}V^{\mu}(x)+\mathrm{i}\theta\theta\bar{\theta}\bar{\lambda}(x)-\mathrm{i}\bar{\theta}\bar{\theta}\theta\lambda(x)+\frac{1}{2}\theta\theta\bar{\theta}\bar{\theta}D(x) \end{aligned} \tag{11.2.29}$$

where, $\Box = \partial_{\mu}\partial^{\mu}$, and C, M, N, D are real scalar fields (spin zero), χ, λ are Weyl spinor fields, V^{μ} is vector field. Obviously

$$V^{+} = V \tag{11.2.30}$$

In Eq. (11.2.16), taking $S(y,\alpha,\bar{\alpha})$ as an infinitesimal SUSY transformation $S(0,\varepsilon,\bar{\varepsilon})$ we can derive

$$\delta\phi(x,\theta,\bar{\theta}) = \left[\varepsilon\frac{\partial}{\partial\theta}+\bar{\varepsilon}\frac{\partial}{\partial\bar{\theta}}-\mathrm{i}\left(\varepsilon\sigma_{\mu}\bar{\theta}-\theta\sigma_{\mu}\bar{\varepsilon}\right)\partial^{\mu}\right]\phi \tag{11.2.31}$$

So, under the infinitesimal SUSY transformation of $S(0,\varepsilon,\bar{\varepsilon})$

$$\delta V = \left[\varepsilon\frac{\partial}{\partial\theta}+\bar{\varepsilon}\frac{\partial}{\partial\bar{\theta}}-\mathrm{i}(\varepsilon\sigma_{\mu}\bar{\theta}-\theta\sigma_{\mu}\bar{\varepsilon})\partial^{\mu}\right]V \tag{11.2.32}$$

We can see that

$$\begin{cases} \delta C = \mathrm{i}\varepsilon\chi-\mathrm{i}\bar{\varepsilon}\bar{\chi},\ \delta\chi = \varepsilon\left(M+\mathrm{i}N\right)+\sigma^{\mu}\bar{\varepsilon}\left(\partial_{\mu}C+\mathrm{i}V_{\mu}\right) \\ \delta\bar{\chi} = \bar{\varepsilon}(M-\mathrm{i}N)+\varepsilon\sigma^{\mu}\left(\partial_{\mu}C-\mathrm{i}V_{\mu}\right), \\ \delta N = \varepsilon\left(\mathrm{i}\lambda-\sigma_{\mu}\partial^{\mu}\bar{\chi}\right)+\bar{\varepsilon}\left(-\mathrm{i}\bar{\lambda}+\bar{\sigma}_{\mu}\partial^{\mu}\chi\right), \\ \delta M = \varepsilon\left(\lambda+\mathrm{i}\sigma_{\mu}\partial^{\mu}\bar{\chi}\right)+\bar{\varepsilon}\left(\bar{\lambda}+\mathrm{i}\bar{\sigma}_{\mu}\partial^{\mu}\chi\right), \\ \delta V_{\mu} = \varepsilon\partial_{\mu}\chi+\bar{\varepsilon}\partial_{\mu}\bar{\chi}+\mathrm{i}\varepsilon\sigma_{\mu}\bar{\lambda}+\mathrm{i}\bar{\varepsilon}\bar{\sigma}_{\mu}\lambda, \\ \delta\lambda = \varepsilon\sigma^{\mu\nu}V_{\mu\nu}+\varepsilon D, \delta\bar{\lambda} = \bar{\varepsilon}\bar{\sigma}^{\mu\nu}V_{\mu\nu}+\bar{\varepsilon}D \\ \delta D = -\varepsilon\sigma^{\mu}\partial_{\mu}\bar{\lambda}+\bar{\varepsilon}\sigma^{\mu}\partial_{\mu}\lambda \end{cases} \tag{11.2.33}$$

where

$$\begin{cases} \sigma^\mu = \left(1, \vec{\sigma}\right), \bar{\sigma}^\mu = (1, -\vec{\sigma}), V_{\mu\nu} = \partial_\mu V_\nu - \partial_\nu V_\mu \\ \sigma^{\mu\nu} = \dfrac{1}{4}\left(\sigma^\mu \bar{\sigma}^\nu - \sigma^\nu \bar{\sigma}^\mu\right), \quad \bar{\sigma}^{\mu\nu} = \dfrac{1}{4}\left(\bar{\sigma}^\mu \sigma^\nu - \bar{\sigma}^\nu \sigma^\mu\right) \end{cases} \tag{11.2.34}$$

From Eq. (11.2.33) we see that the change of the D term δD is also a total derivative.

11.2.3.2. *Constructing lagrangian*

Because Eqs. (11.2.28) and (11.2.23) tell us that δF and δD are all the space–time total derivatives, assuming that the fields turn to zero at infinity, then $\int \mathrm{d}^4x \delta F$ and $\int \mathrm{d}^4x \delta D$ all turn to zero. Therefore $\int \mathrm{d}^4xF$ and $\int \mathrm{d}^4xD$ are invariant under SUSY transformation. Thus we can use this point to construct Lagrangian.

In fact, Lagrange density $\mathscr{L}$ can be written as

$$\mathscr{L} = \mathscr{L}_F + \mathscr{L}_D \tag{11.2.35}$$

Note that

$$\begin{cases} \displaystyle\int \mathrm{d}\theta = 0, \quad \int \mathrm{d}\theta \cdot \theta = 1, \quad \int \mathrm{d}^2\theta \cdot \theta\theta = 1 \\ \displaystyle\int \mathrm{d}\bar{\theta} = 0, \quad \int \mathrm{d}\bar{\theta} \cdot \bar{\theta} = 1, \quad \int \mathrm{d}^2\bar{\theta} \cdot \bar{\theta}\bar{\theta} = 1 \end{cases} \tag{11.2.36}$$

In $\theta\theta F$, $\theta\theta\bar{\theta}\bar{\theta}D$ terms F and D can be extracted by integration with respect to θ, $\bar{\theta}$. So the Lagrangian

$$L = \int \mathrm{d}^4x \left[\int \mathrm{d}^2\theta \mathrm{d}^2\bar{\theta} \mathscr{L}_D + \int \mathrm{d}^2\theta \mathscr{L}_F + \text{h.c.}\right] \tag{11.2.37}$$

To construct Lagrangian we need bilinear superfields, i.e. $\phi\phi^+$-like terms.

For a left-handed chiral superfield ϕ with $\bar{D}\phi = 0$, the ϕ^2, ϕ^n all satisfy $\bar{D}\phi^2 = \bar{D}\phi^n = 0$, i.e. ϕ^2, $\phi^n (n \geq 2)$ are all left-handed chiral superfields. From Eq. (11.2.26), we also see that, ϕ^+ and $(\phi^+)^n (n \geq 2)$ are all right-handed chiral superfields (satisfying $D_{\mathrm{R}}\phi^+ = 0$, $D_{\mathrm{R}}(\phi^+)^n = 0$). Therefore, the product of left-handed chiral superfields is always left-handed chiral superfield, the product of right-handed chiral superfields is always right-handed chiral superfield.

Equation (11.2.26) leads to

$$\phi_{\mathrm{L}}^{2}\left(x,\theta\right)=\varphi^{2}\left(x\right)+2\varphi\left(x\right)\theta^{\alpha}\psi_{\alpha}(x)+(\theta\theta)[2\varphi F-\frac{1}{2}\psi^{\alpha}\psi_{\alpha} \tag{11.2.38}$$

$$\int \mathrm{d}^{2}\theta\varphi_{\mathrm{L}}^{2}(x,\theta)=2\varphi F-\frac{1}{2}\psi^{\alpha}\psi_{\alpha} \tag{11.2.39}$$

We can see that the F term in lagrangian comes always from product of left-handed chiral superfields (it can have $\theta\theta$ term). It can be proved that, ϕ_{L}, ϕ_{R} have the following relation [16]:

$$\phi_{\mathrm{L}}(x_{\mu}+\mathrm{i}\theta\sigma_{\mu}\bar{\theta},\theta,\bar{\theta})=\phi_{\mathrm{R}}(x_{\mu}-\mathrm{i}\theta\sigma_{\mu}\bar{\theta},\theta,\bar{\theta}) \tag{11.2.40}$$

For constructing Lagrangian we need bilinear form of $\phi\phi^{+}$. But construction of $\phi\phi^{+}$ needs ϕ and ϕ^{+} to be in the same left-handed representation. Because $(\phi_{\mathrm{L}})^{+}$ is in right-handed representation, so we need to transform $(\phi_{\mathrm{L}})^{+}$ into left-handed representation by use of Eq. (11.2.40).

Taking conjugation to Eq. (11.2.26), we get right-handed chiral superfield

$$\phi_{(\mathrm{R})}^{+}(x,\bar{\theta})=\varphi^{*}+(\bar{\theta}\bar{\psi})+(\bar{\theta}\bar{\theta})F^{*} \tag{11.2.41}$$

Using Eq. (11.2.40) we can transform Eq. (11.2.41) into left-handed representation

$$\phi_{(\mathrm{L})}^{+}(x_{\mu},\theta,\bar{\theta})=\phi_{(\mathrm{R})}^{+}(x_{\mu}-2\mathrm{i}(\theta\sigma_{\mu}\bar{\theta}),\theta,\bar{\theta}) \tag{11.2.42}$$

Substituting Eq. (11.2.42) into Eq. (12.2.41) and expending it at the point x_{μ} we have

$$\begin{aligned}\phi_{(\mathrm{L})}^{+}(x_{\mu},\theta,\bar{\theta})&=\varphi^{*}-2\mathrm{i}(\theta\sigma_{\mu}\bar{\theta})\partial^{\mu}\varphi^{*}-2(\theta\sigma_{\mu}\bar{\theta})(\theta\sigma_{\nu}\bar{\theta})\partial^{\mu}\partial^{\nu}\varphi^{*}\\&\quad+(\bar{\theta}\bar{\psi})-2\mathrm{i}(\theta\sigma_{\mu}\bar{\theta})\partial^{\mu}(\bar{\theta}\bar{\psi})+(\bar{\theta}\bar{\theta})F^{*}\end{aligned} \tag{11.2.43}$$

So the product of Eq. (11.2.26) and Eq. (11.2.43) leads to

$$\begin{aligned}\phi(x,\theta)\phi^{+}(x,\theta,\bar{\theta})&=\phi(x,\theta)\phi^{+}(x-2\mathrm{i}(\theta\sigma\bar{\theta}),\bar{\theta})\\&=\phi(x,\theta)\left[\mathrm{e}^{-2\mathrm{i}\theta\sigma_{\mu}}\bar{\theta}\partial^{\mu}\right]\phi^{+}(x,\bar{\theta})\end{aligned} \tag{11.2.44}$$

In Eq. (11.2.44) $\phi(x,\theta)$ is just the field of Eq. (11.2.26), $\phi^{+}(x,\bar{\theta})$ is just the field of Eq. (11.2.41). Equation (11.2.44) is a vector superfield, from it we can compute out

$$(\phi\phi^+)_D = FF^* - \varphi\Box\varphi^* + \frac{1}{2}\psi^\beta\sigma^\mu_{\beta\dot{\alpha}}\partial_\mu\bar{\psi}^{\dot{\alpha}} \tag{11.2.45}$$

According to Eq. (11.2.22), $\int \mathrm{d}^4x(\phi\phi^+)_D$ is SUSY invariant. Equation (11.2.45) can obviously represents the kinetic energy terms of the scalar field φ and the Weyl spinor field ψ. Their mass and interaction terms can be given from the F term. The F term $\mathscr{L}_F$ is also called super potential $W(\phi)$

$$W(\phi) = \mathscr{L}_F = m\phi^2 + \lambda\phi^3 \tag{11.2.46}$$

The Lagrangian can be written as

$$\mathscr{L} = (\phi\phi^+)_D + m(\phi^2)_F + \lambda(\phi^3)_F + \text{h.c.} \tag{11.2.47}$$

Neglecting the total derivative terms, Eq. (11.2.47) can be rewritten by use of the component fields as

$$\begin{aligned}\mathscr{L} = (\partial_\mu\varphi)\left(\partial^\mu\varphi^*\right) + \frac{1}{2}\mathrm{i}\psi\sigma_\mu\partial^\mu\bar{\psi} + FF^* + m\left(2\varphi F - \frac{1}{2}(\psi\psi) + \text{h.c.}\right)\\ + \lambda\left(3\varphi^2F - \frac{3}{2}\varphi(\psi\psi) + \text{h.c.}\right)\end{aligned} \tag{11.2.48}$$

From this, we can derive the equation of motion

$$\begin{cases}\partial^\mu\partial_\mu\phi = 2mF^* - \dfrac{3}{2}\lambda\left(\bar{\psi}\bar{\psi}\right) + 6\lambda\varphi^*F^*\\ \mathrm{i}\sigma_\mu\partial^\mu\bar{\psi} = 2m\psi + 3\lambda\varphi\psi\\ F^* + 2m\varphi + 3\lambda\varphi^2 = 0\end{cases} \tag{11.2.49}$$

From the third formula of Eq. (11.2.49) we can eliminate the auxiliary field F, i.e. substituting $F = -2m\varphi^* - 2\lambda\varphi^{*2}$ into Eq. (11.2.48) we have

$$\begin{aligned}\mathscr{L} = \left|\partial_\mu\varphi\right|^2 + \frac{1}{2}\mathrm{i}\psi\sigma_\mu\partial^\mu\bar{\psi} + \left|2m\varphi + 3\lambda\varphi^2\right|^2\\ - m\left[2\varphi(2m\varphi^* + 3\lambda\varphi^{*2}) + \frac{1}{2}(\psi\psi) + \text{h.c.}\right]\\ - \lambda\left[3\varphi^2\left(2m\varphi^* + 3\lambda\varphi^{*2}\right) + \frac{3}{2}\varphi\left(\psi\psi\right) + \text{h.c.}\right]\end{aligned} \tag{11.2.50}$$

Equation (11.2.50) is just the lagrangian for describing spin-0 and spin-1/2 fields.

In the Lagrangian the part which does not include fermions and derivatives can be defined as scalar potential, and expressed as $V(\varphi)$, then $V(\varphi)$ can be obtained from the superpotential [17]

$$V(\varphi) = \left|\partial W(\varphi)/\partial\varphi\right|^2 \tag{11.2.51}$$

where $W(\varphi)$ is obtained by replacing the chiral superfield ϕ in the superpotential $W(\phi) \equiv \mathscr{L}_F$ (see Eq. (11.2.46)) by its scalar field φ. Obviously

$$V(\varphi) = \left|\partial W(\varphi)/\partial\varphi\right|^2 = \left|2m\varphi + 3\lambda\varphi^2\right|^2 \tag{11.2.52}$$

Expressing the h.c.term of Eq. (11.2.50) explicitly and making arrangement, we obtain

$$\begin{aligned}\mathscr{L} = {}& \left|\partial_\mu\varphi\right|^2 + \frac{1}{2}\mathrm{i}\psi\sigma_\mu\partial^\mu\bar{\psi} - \frac{1}{2}m\left(\psi\psi + \bar{\psi}\bar{\psi}\right) \\ & - \frac{3}{2}\lambda\left(\varphi\psi\psi + \varphi^*\bar{\psi}\bar{\psi}\right) - \left|2m\varphi + 3\lambda\varphi^2\right|^2 \end{aligned} \tag{11.2.53}$$

In the following, for supersymmetric Yang–Mills theory which includes gauge fields, we only give the results without derivation. The readers who are interested in can go to the original references.

Let us see Abelian gauge fields first. In that case we need to introduce the vector superfield $V(x,\theta,\bar{\theta})$ of Eq. (11.2.29). Its vector component V_μ transforms in gauge transformation of Abelian case as

$$V_\mu \to V_\mu + \partial_\mu\varphi \tag{11.2.54}$$

φ is real scalar field. In order to put on SUSY, φ must belong to a supermultiplet. For Wess–Zumino gauge [18], the vector superfield V transforms according to

$$V \to V + \mathrm{i}(\Lambda - \Lambda^+) \tag{11.2.55}$$

When Λ is a chiral superfield, it includes a complex scalar field η and a Weyl spinor field, while φ in Eq. (11.2.54) is here as $\varphi = (\eta + \eta^*)/2$. For the Wess–Zumino gauge C, χ, M, N in the vector superfield of Eq. (11.2.29) all can be gauged away, but λ, D are gauge invariant, while V_μ transforms still according to Eq. (11.2.54). In that case, the vector superfield is simplified as

$$V_{\mathrm{WZ}}(x,\theta,\bar{\theta}) = -\theta\sigma_\mu\bar{\theta}V^\mu + \mathrm{i}\theta\theta\bar{\theta}\bar{\lambda} - \mathrm{i}\bar{\theta}\bar{\theta}\theta\lambda + \frac{1}{2}\theta\theta\bar{\theta}\bar{\theta}D \tag{11.2.56}$$

Wess–Zumino gauge is a unitary gauge, it is suited to constructing Lagrangian but not good to do perturbation calculation.

For extending to the case of non-Abelian gauge fields, we introduce a chiral superfield

$$W_\alpha = (\bar{D}\bar{D})D_\alpha V \tag{11.2.57}$$

where $\bar{D}_{\dot{\alpha}}$, D_α are given in Eq. (11.2.22). In left-handed (L) representation,

$$\begin{aligned} W_\alpha(x,\theta) = 4\mathrm{i}\lambda_\alpha - 4\theta_\alpha D + 4\mathrm{i}\theta^\beta \sigma_{\nu\alpha\dot{\beta}} \sigma^{\dot{\beta}}_{\mu\beta} (\partial^\mu V^\nu - \partial^\nu V^\mu) \\ - 4(\theta\theta)\sigma_{\nu\alpha\dot{\beta}} \partial^\mu \bar{\lambda}^{\dot{\beta}} \end{aligned} \tag{11.2.58}$$

where D, λ, $V_{\mu\nu} = \partial_\mu V_\nu - \partial_\nu V_\mu \equiv F_{\mu\nu}$ are all gauge invariant. So, the F term of the chiral field $W^\alpha W_\alpha$ is not only gauge invariant but also SUSY invariant,

$$\begin{aligned} \frac{1}{32}\left(W^\alpha W_\alpha\right)_F = -\frac{1}{4} V_{\mu\nu} V^{\mu\nu} - \frac{1}{2}\mathrm{i}\lambda^\alpha \sigma_{\mu\alpha\dot{\gamma}} \partial^\mu \bar{\lambda}^{\dot{\gamma}} \\ - \frac{1}{2}\mathrm{i}\sigma^\alpha_{\mu\dot{\beta}} \left(\partial^\mu \bar{\lambda}^{\dot{\beta}}\right) \lambda_\alpha + \frac{1}{2} D^2 \end{aligned} \tag{11.2.59}$$

This is just the Lagrangian of the supersymmetrized pure Abel Yang–Mills theory. The auxiliary field D can be eliminated by use of the equation of motion, while λ is just the superpartner of V^μ gaugino. The minimal coupling with the matter field $\phi = (\varphi, \psi, F)$ is

$$\begin{aligned} \left(\phi^+ \exp(2gV)\,\phi\right)_D = |D_\mu \varphi|^2 - \frac{1}{2}\mathrm{i}\left[\bar{\psi}^{\dot{\beta}} \sigma_{\mu\alpha\dot{\beta}} D^\mu \psi^\alpha\right] \\ + g\varphi^* D\varphi + \mathrm{i}g\left[\varphi^*(\lambda\psi) - \left(\bar{\lambda}\bar{\psi}\right)\varphi\right] + |F|^2 \end{aligned} \tag{11.2.60}$$

where g is the coupling constant of V^μ with matter field, $D_\mu = \partial_\mu + \mathrm{i}gV_\mu$ is the covariant derivative under Wess–Zumino gauge.

For extending to non-Abelian case [19] we can put

$$V = V^a T^a, \quad \Lambda = \Lambda^a T^a \tag{11.2.61}$$

where T^a are the generators of the non-Abelian group. In that case W_α can be extended to

$$W_\alpha = \bar{D}\bar{D}\left[\exp(-gV)\, D_\alpha \exp(gV)\right] \tag{11.2.62}$$

where g is the gauge coupling constant. The gauge transformations are

$$\begin{cases} \mathrm{e}^{gV} \to \mathrm{e}^{-\mathrm{i}g\Lambda^+} \mathrm{e}^{gV} \mathrm{e}^{\mathrm{i}g\Lambda} \\ W_\alpha \to \mathrm{e}^{-\mathrm{i}g\Lambda} W_\alpha \mathrm{e}^{\mathrm{i}g\Lambda} \end{cases} \tag{11.2.63}$$

The Lagrangian density is

$$\begin{aligned}\mathscr{L} &= \frac{1}{32}g^2(W^\alpha W_\alpha)_{\mathrm{F}} + (\phi^+ \mathrm{e}^{2gV}\phi)_{\mathrm{D}} \\ &= -\frac{1}{4}G_{\mu\nu}G^{\mu\nu} + \frac{1}{2}D^2 - \frac{1}{2}\mathrm{i}\{\lambda^\alpha \sigma_{\mu\alpha\dot{\beta}}(\partial^\mu \bar{\lambda}^{\dot{\beta}} + \mathrm{i}g[V^\mu, \bar{\lambda}^{\dot{\beta}}]) \\ &\quad - (\partial^\mu \bar{\lambda}^{\dot{\beta}} + \mathrm{i}g[V^\mu, \bar{\lambda}^{\dot{\beta}}])\sigma_{\mu\alpha\dot{\beta}}\lambda^\alpha\} + (\phi^+ \mathrm{e}^{2gV}\phi)_{\mathrm{D}} \qquad (11.2.64)\end{aligned}$$

The last term is the minimal coupling with matter field in Eq. (11.2.60).

Above, we only gave the calculation method for the Lagrangian in a very simply way. The readers who want to work in SUSY field should refer to the relative original references.

11.2.3.3. *Minimal supersymmetric standard model*

If we supersymmetrize the standard model of the $\mathrm{SU}(3)_{\mathrm{C}} \times \mathrm{SU}(2) \times \mathrm{U}(1)$ gauge theory, then we get the minimal supersymmetric standard model (MSSM).

In order to supersymmetrize the standard model, we need to introduce a series of superfields whose supermultiplets include the standard model particles and their superpartners. These superfields (supermultiplets) are listed in Tables 11.1 and 11.2, where $i = 1, 2, 3$ are generation or family indices, $\alpha = 1, 2$ are the weak isospin indices, and the $\mathrm{SU}(3)_{\mathrm{C}}$ color indices are omitted.

Table 11.1. Chiral superfields in MSSM.

Superfield		Spin 1/2	Spin 0	$\mathrm{SU}(3)_{\mathrm{C}}, \mathrm{SU}(2)_L, \mathrm{U}(1)_Y$
Quarks, squarks (×3 families)	Q_i^α	$(u_{i\mathrm{L}}, d_{i\mathrm{L}})$	$\left(\tilde{u}_{i\mathrm{L}}, \tilde{d}_{i\mathrm{L}}\right)$	$(3, 2, 1/6)$
	$\bar{U}_i$	$u_{i\mathrm{R}}^+$	$\tilde{u}_{i\mathrm{R}}^*$	$(\bar{3}, 1, -2/3)$
	$\bar{D}_i$	$d_{i\mathrm{R}}^+$	$\tilde{d}_{i\mathrm{R}}^*$	$(\bar{3}, 1, 1/3)$
Leptons, sleptons (×3 families)	L_i^α	$(\nu_i, e_{i\mathrm{L}})$	$(\widetilde{\nu}_i, \tilde{e}_{i\mathrm{L}})$	$(1, 2, -1/2)$
	$\bar{E}_i$	$e_{i\mathrm{R}}^+$	$\tilde{e}_{i\mathrm{R}}^*$	$(1, 1, 1)$
Higgs, higgsinos	H_u^α	$\left(\tilde{H}_u^+, \tilde{H}_u^0\right)$	(H_u^+, H_u^0)	$(1, 2, +1/2)$
	H_d^α	$\left(\tilde{H}_d^0, \tilde{H}_d^-\right)$	$\left(H_d^0, H_d^-\right)$	$(1, 2, -1/2)$

Note: spin-0 fields are complex scalar fields, spin-1/2 fields are left-handed two component Weyl spinors.

Table 11.2. Gauge super multiplets in MSSM

	Spin 1	Spin 1/2	$\mathrm{SU}(3)_\mathrm{C}, \mathrm{SU}(2)_L, \mathrm{U}(1)_Y$
Gluon, gluino	g	$\tilde{g}$	$(8, 1, 0)$
W-bosons, winos	$W^\pm, W^0$	$\tilde{W}^\pm, \tilde{W}^0$	$(1, 3, 0)$
B-boson, bino	B^0	$\tilde{B}^0$	$(1, 1, 0)$

W^0 and B^0 are not weak interaction eigenstates, after mixing they become γ and Z^0, the corresponding superpartners are called photino and zino, respectively. In Tables 11.1 and 11.2, the superpartners are often mixed, the mixed mass eigenstates of $\tilde{W}^\pm$ and $\tilde{H}^\pm$ are collectively called charginos and expressed by $\tilde{\chi}^\pm_{1,2}$. The mixed mass eigenstates of the neutral superpartners are collectively called neutralinos and expressed by $\tilde{\chi}^0_i$.

When the above superfields are ready, we can construct the Lagrangian of MSSM using the method we mentioned before. We shall not discuss it in detail here, but we give the super-potential W_MSSM, i.e.

$$W_\mathrm{MSSM} = \bar{U} y_u Q H_u + \bar{D} y_d Q H_d + \bar{E} y_e L H_d + \mu H_u H_d \tag{11.2.65}$$

In the above equation, we compress all kinds of indices y_u, y_d, y_e are Yukawa couplings of 3×3 matrices. $\bar{U} y_u Q H_u = \bar{U}^a_i y_{uij} Q^\alpha_{ja} H_{u\beta} \varepsilon^{\alpha\beta}$, $a = 1, 2, 3$ are $\mathrm{SU}(3)_\mathrm{C}$ color indices, α, β are weak isospin indices, $i, j = 1, 2, 3$ are family indices.

If we integrate Eq. (11.2.65) $\int \mathrm{d}^2\theta W_\mathrm{MSSM}$, then we get the Yukawa coupling of the F term

$$L_Y = \lambda_{ij} \bar{u}_{i\mathrm{L}} H^0_u u_{j\mathrm{R}} + \lambda'_{ij} \bar{d}_{i\mathrm{L}} H^0_d d_{j\mathrm{R}} + \lambda''_{ij} \bar{E}_{i\mathrm{L}} H^0_d E_{j\mathrm{R}} \tag{11.2.66}$$

So we see that if we want to give up-quark and down-quark masses simultaneously, we must introduce two sets of Higgs fields. Of course, only in that case of two sets of Higgs fields the loop diagram divergences can be eliminated.

The following two SUSY invariant terms which are allowed to exist and break lepton number and baryon number conservation, respectively, are not included in superpotential of Eq. (11.2.65).

$$W_{\Delta L=1} = \frac{1}{2} \lambda^{ijk} L_i L_j \bar{E}_k + \lambda^{ijk} L_i Q_j \bar{D}_k + \mu'^i L_i H_u \tag{11.2.67}$$

$$W_{\Delta B=1} = \frac{1}{2} \lambda''^{ijk} \bar{U}_i \bar{D}_j \bar{D}_k \tag{11.2.68}$$

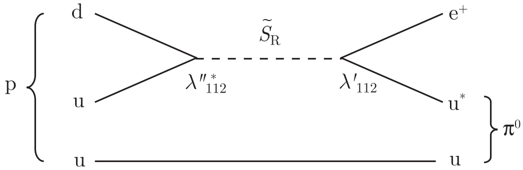


Fig. 11.4. Branching ratio of the proton decay $p \to e^+\pi^0$ caused by the exchange of s-quark superpartner will be larger than the experimental upper bound by several orders of magnitude.

When $W_{\Delta L=1}$ breaks the total lepton number conservation, this term is not allowed by experiments, while $W_{\Delta L=1}$ and $W_{\Delta B=1}$ make the proton decay too fast so violates the experimental constrain. The proton decay diagram is shown in Fig. 11.4.

In order to prevent troubles caused by the two terms $W_{\Delta L=1}$ and $W_{\Delta B=1}$ in Eqs. (11.2.57) and (11.2.68), we can introduce R-symmetry [20] in MSSM models, and we can also introduce discrete productive R-parity to forbid terms of $W_{\Delta L=1}$ and $W_{\Delta B=1}$. This is the most economic method in which all the standard model particles possess R-parity $R_p = 1$, while their superpartners $R_p = -1$. The R-parity conservation demand that experimentally produced superpartners in the final state must be in pair. Therefore, if R_p conservation is exact, then there must be three phenomenological conclusions:

(1) Lightest supersymmetric particle (LSP) should be absolutely stable, because its $R_p = -1$ so cannot decay to any final states. If LSP is neutral (electric charge is zero), so it can only has very weak interactions and becomes a non-baryonic dark matter condidate. It is just the hot particle searched experimentally nowadays.
(2) Each superpartner (except LSP) can only decay into final states with odd number of LSP (in general only one).
(3) At colliding machines, superpartners can only be produced in pair.

MSSM has many free parameters. It is counted carefully that [21] in its Lagrangian there are totally 105 free parameters which include masses, mixing angles, independent phases, etc. This makes MSSM to have little predictive power.

There is another important problem, that is the problem of SUSY breaking. We know that SUSY must be broken at low energy. Usually, we add soft SUSY breaking term artificially in MSSM. These soft supersymmetry breaking terms can only appear in the type of mass term so as to not cause trouble. For example

$$\begin{aligned} L_{\text{soft}}^{\text{MSSM}} = &-\frac{1}{2}\left(M_3\tilde{g}\tilde{g} + M_2\tilde{W}\tilde{W} + M_1\tilde{B}\tilde{B} + \text{h.c.}\right) \\ &-\left(\tilde{\bar{\bar{U}}}a_u\tilde{Q}H_u - \tilde{\bar{\bar{D}}}a_d\tilde{Q}H_d - \tilde{\bar{\bar{E}}}a_e\tilde{L}H_d + \text{h.c.}\right) \\ &-\tilde{Q}^+m_{\tilde{Q}}^2\tilde{Q} - \tilde{L}^+m_{\tilde{L}}^2\tilde{L} - \tilde{\bar{\bar{U}}}m_{\tilde{\bar{\bar{U}}}}^2\tilde{\bar{\bar{U}}}^+ - \tilde{\bar{\bar{D}}}m_{\tilde{\bar{\bar{D}}}}^2\tilde{\bar{\bar{D}}}^+ \\ &-\tilde{\bar{\bar{E}}}m_{\tilde{\bar{\bar{e}}}}^2\tilde{\bar{\bar{E}}}^+ - m_{H_u}^2H_u^*H_u - m_{H_d}^2H_d^*H_d \\ &-\left(bH_uH_d + \text{h.c.}\right) \end{aligned} \tag{11.2.69}$$

where the operators with "$\sim$" on their head are corresponding superpartner fields, a_u, a_d, a_e are all 3×3 complex matrices, and all mass parameters $\sim m_{\text{soft}} \approx< 1\text{TeV}$.

Obviously, artificially introduced soft breaking terms will increase more free mass parameters. How to test MSSM is one of the frontier of the theoretical and experimental researches.

11.3. Superstring Theory and Brane Theory

Superstring theory became fashionable in particle physics community in the 1980s of the 20 century. In superstring theory there are no free parameters, it not only unifies the four interactions (strong, weak, electromagnetic, and gravitational interactions), but also is renormalizable. All physically observed quantities can be computed out theoretically, so string theory is also called "theory of every thing". Superstring theory takes different particles as different string vibration modes. Particles are no longer described as geometric points but as 1-dimensional extended objects (strings). Superstrings are divided into open strings and closed strings, while the string theory more closed to realistic physics is the heterotic string in 10-dimensional space. Although superstring theory is very attractive, but it has a basic difficulty, i.e. when the theory compacts from 10-dimensional space into the real world of $3+1$-dimensional space–time, there will be infinite number of vacuum states, and we have no good way to choose out the vacuum state of our real world.

The readers who are interested in superstrings can refer to a set of very good books [22].

The brane theory is a theory constructed on an 11-dimensional brane. Some special brane theories can deteriorate into 10-dimensional superstring theories after compacting 1-dimensional space. Brane theories are also fashionable ones to unify the four types of interactions. The readers who are interested in can refer to a popular book [23].

References

1. Georgi H, Glashow S L. *Phys. Rev. Lett.*, 1974, 32: 438.
2. Itzykson C. *Rev. Mod. Phys.*, 1966, V38: 95.
3. Buras A J, *et al. Nucl. Phys.*, 1978, B135: 66.
4. Georgi H, Jarlskog C. Havard Univ. Preprint HUTP-79/A026.
5. PDG (Particle Data Group). *Phys. Rev.*, 2012, D86: 010001.
6. Marciano W. *Talk presented at the Forth Workshop on Grand Unification. Univ.* Pennsylvania, April, 1983: 21–23.
7. Makoto Miura (Super–Kamiokande Collab.) *ICHEP* 2010.
8. Georgi H. *Particles and Fields*, ed. Carlson C. *Proc. APS Div. Particles and Fields*, 1975: 575; Fritzsch H, Minkowski P. *Ann. Phys.*, 1975, 93: 193.
9. Du D S (Tu T S), *et al. Sci. Sinica A*, 1981, VXXIV:1366.
10. Ma Z Q, *et al. Phys. Lett.*, 1981, 100B: 399; Ma Z Q, *et al. Phys. Lett.*, 1981, 99B: 107.
11. Ma Z Q, *et al. High Ener. Phys. Nucl. Phys.,* 1981, V5: 664.
12. Dong F X, *et al. Ann. Phys.*, 1983, V145: 1.
13. Kazakov D I. hep-ph/0012288V2.
14. Coleman S, Mandula *J. Phys. Rev.*, 1967, 159: 1251.
15. Grisaru M T, Siegel W, Rocek M. *Nucl. Phys.*, 1979, B159: 429.
16. Nilles N. P. *Phys. Rep.*, 1984, 110: 1; Haber H E, Kane G L. *Phys. Rep.*, 1985, 117: 75; Martin S P. hep-ph/9709356.
17. Witten E. *Nucl. Phys.*, 1981, B188: 513.
18. Wess J, Zumino B. *Nucl. Phys.*, 1974, B78: 1.
19. Ferrara S, Zumino B. *Nucl. Phys.*, 1975, B79: 207; Salam A, Strathdee *J. Phys. Lett.*, 1974, 51B: 353.
20. Fayet P. *Nucl. Phys.*, 1975, B90: 104; Salam A, Strathdee *J. Nucl. Phys.*, 1975, B87: 85.
21. Dimopoulos S, Sutter D. *Nucl. Phys.*, 1995, B452: 496.
22. Green M B, Schwarz J H, Witten E. *String Theory*. Cambridge University Press,1986, 1: 2.
23. Clifford V. *Johnson "D-branes"*. Cambridge University Press, 2003.

Chapter 12

Particle Physics and Cosmology

The standard model of particle physics is an SU(3) × SU(2) × U(1) gauge theory which describes various particles and their interactions in the microscopic world. After being put into countless experimental tests deep to a distance scale of $\mathcal{O}(10^{-17})$ cm, the standard model has proved to be a correct quantum field theory. In cosmology there is also a standard model — the Big Bang ΛCDM model, which can describe the macroscopic world very well. The basic conjecture of this model is that the Universe expands quickly after the Big Bang, and its evolution is largely controlled by dark energy (or equivalently, Einstein's cosmological constant Λ). The dark matter is dominated by the cold dark matter (CDM), and the present Universe is found to be essentially flat. The correctness of such a standard model of cosmology has been verified by many astronomical and cosmological observations, up to a distance scale of $\mathcal{O}(10^{28})$ cm (i.e. larger than ten billion light years).

It is obvious that the distance scales associated with the validity of the Standard Model and the ΛCDM model can span 45 orders of magnitude. Does one of them have something to do with the other in this case? The answer is affirmative. What role does the particle physics play in describing the evolution of the Universe? In this chapter, we shall elucidate these two questions.

Einstein's General Relativity plays the central role in describing the origin and evolution of the Universe. It has been experimentally verified to a high degree of accuracy [1]. In particular, the Laser Interferometer Gravitational-Wave Observatory (LIGO) in the United States has recently discovered gravitational waves emitted from the merging of two black holes about 1.3 billion years ago [2], and the Event Horizon Telescope (EHT)

Collaboration has published the first image of a black hole (whose mass is $(6.5 \pm 0.7)\ M_\odot$ and whose distance from the Earth is about 55 million light years) at the center of galaxy Messier 87 [42]. Such observations have verified all the six well-known predictions of General Relativity:

- The gravitational redshift of a light wave as it moves away from a gravitational field (i.e. its wavelength becomes larger).
- The gravitational deflection of light in a gravitational field.
- The anomalous perihelion precession of the planet Mercury.
- The gravitational time dilation.
- Gravitational waves.
- Black holes.

That is why General Relativity has been widely accepted as the correct theory of gravity. The recent discoveries of gravitational waves are inaugurating a new era in observing the Universe with the help of gravitational waves. In other words, a new science named "gravitational-wave astronomy" is beginning to spring up.

This chapter provides a brief introduction to two main predictions of the standard ΛCDM model: the cosmic microwave background (CMB) radiation and the abundances of light elements in the Universe. Some other cosmological problems, such as the baryon–antibaryon asymmetry, dark matter and dark energy, will also be discussed. For completeness, the inflation theory will be introduced to overcome three typical difficulties associated with the Big Bang model. Let us begin with the Big Bang ΛCDM model.

12.1. Basics of the Big Bang ΛCDM Model

Our Universe is homogeneous and isotropic when it is viewed on a sufficiently large scale. General Relativity predicts that the Universe must expand. As early as in 1929, Hubble has observed various galaxies receding from the Earth — a robust evidence for the expansion of the Universe [3]. But Hubble's observational results cannot be regarded as the most important pilar of the Big Bang cosmology. Gamow, Alpher and Herman assumed that the Universe was initially in a state of very high temperature and very high density [4], and then it started to expand and cool down until now. The high density and high temperature of the early Universe made it possible to trigger the primordial nucleosynthesis. With the help of the Big Bang cosmology, Alpher and Herman predicted today's CMB temperature to be around several kelvins in 1948 [5]. So the key experimental evidence for the correctness of the Big Bang cosmology should be the observations of the primordial abundances of light elements and the CMB radiation of about three kelvins, which make this model distinguishable from the others.

12.1.1. *Hubble's law*

In 1929, Hubble observed that the light emitted from distant galaxies had a redshift when it reached the Earth. Namely,

$$\lambda' = \lambda\,(1+z) \tag{12.1.1}$$

where λ is the wavelength of the light emitted from distant galaxies, λ' is the wavelength of the light arriving at the Earth, and

$$z = \frac{\lambda' - \lambda}{\lambda} = \frac{\nu - \nu'}{\nu'} \tag{12.1.2}$$

with ν and ν' being the corresponding frequencies. Equation (12.1.1) can be explained with the Doppler effect when $z \ll 1$ is satisfied, where z is referred to as the redshift. The magnitude of z depends on the distance between the observer and galaxies (see Sec. 12.1.4). Hubble found that the velocity of galaxies receding from the Earth (v) is proportional to their distance from the Earth (D):

$$v = HD \tag{12.1.3}$$

where H is the Hubble parameter. Note that v, H and D are all dependent on time t. The physical distance $D(t)$ between the Earth and distant galaxies can be expressed as

$$D(t) = R(t) r_{\rm co} \tag{12.1.4}$$

where $r_{\rm co}$ is a time-independent quantity describing the distance between the Earth and galaxies measured in the comoving frame on the Earth (i.e. the frame that expands together with the Universe), and $R(t)$ is a dimensionless scale factor describing the expansion of the Universe.

It is obvious that $v(t) = \dot{R}(t) r_{\rm co}$ holds. As a result,

$$H(t) = \frac{\dot{R}(t)}{R(t)} \tag{12.1.5}$$

So the Hubble parameter $H(t)$ is time-dependent, and its value measured today reads [6]

$$\begin{cases} H_0 = H(t_0) = 100h_0 \text{km s}^{-1} M_{\text{pc}}^{-1} \\ h_0 = 0.673 \pm 0.012 \end{cases} \tag{12.1.6}$$

where the subscript "0" represents the value of the present time.

Note that all the distances observed by Hubble at that time satisfy $z \ll 1$. When z becomes larger and larger, the Hubble's law has to be modified (in this connection we refer the reader to Sec. 12.7 about dark energy).

12.1.2. *The Robertson–Walker metric and the Friedmann equation*

A homogeneous, isotropic and expanding Universe can be described in terms of the Robertson–Walker metric [7, 8]. In the cartesian coordinate system, the metric tensor $g_{\mu\nu}$, which describes the local geometry of space–time, is given by

$$\begin{cases} g_{00} = -1 \\ g_{0j} = g_{j0} = 0 \\ g_{ij} = g_{ji} = R^2(t) \left(\delta_{ij} + \kappa \dfrac{x^i x^j}{1 - \kappa\, r^2} \right) \end{cases} \tag{12.1.7}$$

and

$$\begin{cases} g^{00} = -1 \\ g^{0j} = g^{j0} = 0 \\ g^{jk} = g^{kj} = \dfrac{1}{R^2(t)} (\delta_{jk} - \kappa x^j x^k) \end{cases} \tag{12.1.8}$$

where i and j run over 1, 2 and 3, $\vec{r} = (x^1, x^2, x^3)$ in the comoving coordinate system, $R(t)$ is the scale factor of the expanding Universe, and κ is a constant representing the curvature of the space:

$$\kappa = \begin{cases} +1 & \text{a closed Universe} \\ 0 & \text{a flat Universe} \\ -1 & \text{an open Universe} \end{cases} \tag{12.1.9}$$

If the spherical coordinates $i, j, k = (r, \theta, \varphi)$ are used, the metric tensor can be expressed as

$$\begin{cases} g_{00} = -1 \\ g_{rr} = \dfrac{R^2(t)}{1 - \kappa\, r^2} \\ g_{\theta\theta} = R^2(t) r^2 \\ g_{\varphi\varphi} = R^2(t) r^2 \sin^2\theta \end{cases} \tag{12.1.10}$$

Given the Robertson–Walker metric, one may write out the line element

$$\begin{cases} \mathrm{d}s^2 = -g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = \mathrm{d}t^2 - R^2(t)\left(\dfrac{\mathrm{d}r^2}{1-\kappa r^2} + r^2\mathrm{d}\Omega^2\right) \\ \mathrm{d}\Omega^2 = \mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\varphi^2 \end{cases} \tag{12.1.11}$$

in the spherical coordinate system. Note that $\sqrt{\mathrm{d}s^2}$ is referred to as the proper time, while t is called the cosmological proper time or cosmic time. The latter is the time measured by a comoving observer — the observer stays still in the comoving coordinate system.

The dynamics and evolution of the Universe are determined by three sets of equations: the Einstein field equations, the energy–momentum relation and the equation of state. Given the definition of the Ricci tensor $R_{\mu\nu}$ [6,7], the Einstein field equations can be written as

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R^\lambda_\lambda = -8\pi G_N T_{\mu\nu} + \Lambda g_{\mu\nu} \tag{12.1.12}$$

It is reasonable to assume that the matter content of the Universe is a perfect fluid on scales larger than that of a galaxy, for which the energy–momentum tensor is given by

$$T^{\mu\nu} = p g^{\mu\nu} + (p + \rho) u^\mu u^\nu \tag{12.1.13}$$

with $u^\mu = (1, 0, 0, 0)$ being the four-velocity vector for the isotropic fluid, p being the isotropic pressure and ρ being the energy density.

With the help of the Robertson–Walker metric, one may derive the Friedmann equation from Eqs. (12.1.12) and (12.1.13) [9]:

$$H^2 \equiv \left(\frac{\dot{R}}{R}\right)^2 = \frac{8\pi G_{\mathrm{N}}\rho}{3} - \frac{\kappa}{R^2} + \frac{\Lambda}{3} \tag{12.1.14}$$

and

$$\frac{\ddot{R}}{R} = \frac{\Lambda}{3} - \frac{4\pi G_{\rm N}}{3}(\rho + 3p) \tag{12.1.15}$$

On the other hand, the energy–momentum conservation law leads us to

$$T^{\mu\nu}_{;\nu} = \frac{\partial}{\partial x^{\nu}} T^{\mu\nu} + \Gamma^{\mu}_{\nu\lambda} T^{\lambda\nu} + \Gamma^{\nu}_{\nu\lambda} T^{\mu\lambda} = 0 \tag{12.1.16}$$

from which one obtains

$$\dot{\rho} = -3H(\rho + p) \tag{12.1.17}$$

This formula can also be derived from Eqs. (12.1.14) and (12.1.15).

Equation (12.1.14) shows that the evolution of the Universe will be fully determined by the curvature constant κ if the term $\Lambda/3$ is neglected (i.e. if $\Lambda = 0$ is taken). In this case $\kappa = +1$ corresponds to a closed Universe, which is expected to first expand to its maximum radius with $\dot{R}(t) = 0$ and then recollapse in a finite time. For $\kappa = -1$, the Universe is open and will expand infinitely. Moreover, $\kappa = 0$ corresponds to a flat Universe, which will also expand infinitely. The latest astronomical observations tell us that today's Universe is spatially close to the flat case.

12.1.3. *Definitions of some cosmological parameters*

Let us proceed to define some observable parameters of the Universe. The Hubble parameter H has been defined in Eq. (12.1.5). Taking $\Lambda = 0$ and $\kappa = 0$ in Eq. (12.1.14), one may define the critical energy density

$$\rho_{\rm c} = \frac{3H^2}{8\pi G_{\rm N}} = 1.88 \times 10^{-26} h^2\,{\rm kgm}^{-3} = 1.05 \times 10^{-5} h^2\,{\rm GeV\ cm}^{-3} \tag{12.1.18}$$

where h is called the "scaled Hubble parameter". Of course, h is also time-dependent, and it is related to H via

$$H(t) = 100h(t){\rm km\ s}^{-1} M_{\rm pc}^{-1} \tag{12.1.19}$$

The present-day value of $\rho_{\rm c}(t_0)$ can be obtained by substituting $h(t_0) = 0.673$ into Eq. (12.1.18):

$$\rho_{\rm c}(t_0) \approx 8.5 \times 10^{-27}\,{\rm kg\ m}^{-3} \approx 4.76 \times 10^{-6}\,{\rm GeV\ cm}^{-3}$$

The cosmological density parameter $\Omega_{\rm tot}$ is defined as

$$\Omega_{\rm tot} = \rho/\rho_{\rm c} \tag{12.1.20}$$

Given the definition

$$\rho_\Lambda = \frac{\Lambda}{8\pi G_{\rm N}} \tag{12.1.21}$$

the total energy density of the Universe can be expressed as

$$\rho = \rho_{\rm m} + \rho_{\rm r} + \rho_\Lambda \tag{12.1.22}$$

where $\rho_{\rm m}$ denotes the energy density for non-relativistic and pressureless matter, $\rho_{\rm r}$ is the energy density for relativistic particles (such as photons), and ρ_Λ stands for the vacuum energy density characterized by the cosmological constant Λ. Note that the vacuum energy density is unnecessarily a constant in general, so it can be denoted as $\rho_{\rm V}$. Then the total energy density of the Universe is given by

$$\rho = \rho_{\rm m} + \rho_{\rm r} + \rho_{\rm V} \tag{12.1.23}$$

According to Eq. (12.1.20), one may define the corresponding density parameters $\Omega_{\rm m}$, $\Omega_{\rm r}$ and $\Omega_{\rm V}$. Therefore,

$$\Omega_{\rm tot} = \Omega_{\rm m} + \Omega_{\rm r} + \Omega_{\rm V} \tag{12.1.24}$$

In this case, the Friedmann equation can be rewritten as

$$\begin{aligned}\frac{\kappa}{R(t)^2} &= H(t)^2(\Omega_{\rm m} + \Omega_{\rm r} + \Omega_{\rm V} - 1)\\ &= H(t)^2(\Omega_{\rm tot} - 1)\end{aligned} \tag{12.1.25}$$

The latest observational values of those dimensionless density parameters in Eq. (12.1.25) are listed as follows [1]:

$$\begin{cases}\Omega_{\rm b} = 0.0499(22)\\ \Omega_{\rm cdm} = 0.265(11)\\ \Omega_{\rm m} = \Omega_{\rm cdm} + \Omega_{\rm b} = 0.315^{+0.016}_{-0.017}\\ \Omega_{\rm V} = \Omega_\Lambda = 0.685^{+0.017}_{-0.016}\\ \Omega_\gamma = \rho_\gamma/\rho_{\rm c} = 5.46(19)\times 10^{-5}\\ \Omega_\nu = \rho_\nu/\rho_{\rm c} < 0.0055\\ \Omega_{\rm r} = \Omega_\gamma + \Omega_\nu\\ \Omega_{\rm tot} = \Omega_{\rm m} + \Omega_{\rm r} + \Omega_\Lambda = 0.96^{+0.4}_{-0.5}\end{cases} \tag{12.1.26}$$

Here the density of relativistic particles is the sum of the densities of photons and neutrinos, whose present-day values are very small. One can see that Ω_{tot} is very close to 1, indicating the fact that today's Universe is spatially very flat.

12.1.4. *The redshift and its relation with the physical distance*

Here we follow Ref. [7] to discuss the redshift issues, and refer the reader to Ref. [10] for some similar discussions. Given the comoving coordinate system whose origin is located somewhere on the Earth, we consider the light emitted from a galaxy with the comoving coordinates (r, θ, φ) and traveling to the origin of this comoving frame along the $-\vec{r}$ direction with θ and φ being fixed. The propagation of wave crest of the light satisfies

$$0 = \mathrm{d}t^2 - R^2(t)\frac{\mathrm{d}r^2}{1-\kappa\, r^2} \tag{12.1.27}$$

Let us suppose that the wave crest of the light leaves a galaxy located on $(r_1, \theta_1, \varphi_1)$ at time t_1 and arrives at the origin of the comoving frame on the Earth at time t_0. Then Eq. (12.1.27) allows us to obtain

$$\int_{t_1}^{t_0} \frac{\mathrm{d}t}{R(t)} = \int_0^{r_1} \frac{\mathrm{d}r}{\sqrt{1-\kappa\, r^2}} = \begin{cases} \sin^{-1} r_1, & \kappa = +1 \\ r_1, & \kappa = 0 \\ \sinh^{-1} r_1, & \kappa = -1 \end{cases} \tag{12.1.28}$$

At time t, the physical distance between this galaxy and the observer on the Earth (i.e. the origin of the comoving frame) is given by

$$D(t) = R(t) r_1 \tag{12.1.29}$$

If the second wave crest of the light (next to the first one) leaves the above galaxy at time $t_1 + \delta t_1$ and arrives at the location of the observer at time $t_0 + \delta t_0$, then Eq. (12.1.27) leads us to

$$\int_{t_1+\delta t_1}^{t_0+\delta t_0} \frac{\mathrm{d}t}{R(t)} = \int_0^{r_1} \frac{\mathrm{d}r}{\sqrt{1-\kappa r^2}} \tag{12.1.30}$$

because r_1 is a time-independent comoving coordinate. As a consequence,

$$\int_{t_1+\delta t_1}^{t_0+\delta t_0} \frac{\mathrm{d}t}{R(t)} = \int_{t_1}^{t_0} \frac{\mathrm{d}t}{R(t)} \tag{12.1.31}$$

from which one may easily obtain

$$\int_{t_1}^{t_1+\delta t_1} \frac{\mathrm{d}t}{R(t)} = \int_{t_0}^{t_0+\delta t_0} \frac{\mathrm{d}t}{R(t)} \tag{12.1.32}$$

Note that the distance between two neighboring wave crests is just equal to the wavelength. As for visible light, both δt_1 and δt_0 are expected to be of $\mathcal{O}(10^{-15})$s, and thus $c\delta t_1$ and $c\delta t_0$ happen to lie in the wavelength range of visible light.

Given $c = 1$ in natural units, one is left with $\delta t_1 = \lambda_1$ and $\delta t_0 = \lambda_0$. The smallness of δt assures that the change of $R(t)$ is very small. In view of Eq. (12.1.32), one has

$$\frac{\delta t_0}{R(t_0)} = \frac{\delta t_1}{R(t_1)} \tag{12.1.33}$$

or equivalently

$$\frac{\lambda_0}{R(t_0)} = \frac{\lambda_1}{R(t_1)} \tag{12.1.34}$$

As a result,

$$\frac{\nu_0}{\nu_1} = \frac{R(t_1)}{R(t_0)} \tag{12.1.35}$$

Defining the redshift

$$z \equiv \frac{\lambda_0 - \lambda_1}{\lambda_1} = \frac{\nu_1 - \nu_0}{\nu_0} \tag{12.1.36}$$

where λ_1 denotes the wavelength of the light leaving the galaxy and λ_0 stands for the wavelength of the light arriving at the Earth, we obtain

$$z = \frac{R(t_0)}{R(t_1)} - 1 \tag{12.1.37}$$

When $\lambda_0 > \lambda_1$, $z > 0$ holds. In this case z is called the redshift. If $z < 0$ (i.e. $\lambda_0 < \lambda_1$) holds, then z is referred to as the blueshift. The expansion and recollapse (shrink) of the Universe correspond to the redshifts and blueshifts, respectively. Note that the motion and gravitation of the galaxy are not taken into account in the above discussions about the redshifts, since their effects are not important in most cases.

At time $t = t_0$ the physical distance between the galaxy located on $(r_1, \theta_1, \varphi_1)$ and the observer at the origin of the comoving frame is

$$D(t_0) = R(t_0)r_1 = (z+1)R(t_1)r_1 \tag{12.1.38}$$

So the larger z is, the greater the physical distance of this galaxy will be.

12.1.5. *Some important solutions to the standard* Λ*CDM model*

12.1.5.1. *The solutions based on a specific equation of state*

As mentioned above, the matter content of the Universe can be regarded as a perfect fluid on scales much larger than that of a galaxy. In this assumption, the equation of state of the Universe is given by [11]

$$p_j = w_j \rho_j \tag{12.1.39}$$

where the subscript j represents the radiation field, matter, vacuum and so on; ρ_j denotes the corresponding energy density; p_j is the pressure; and w_j is a constant. To be specific,

$$w_j = \begin{cases} 1 & \text{the massless scalar field} \\ 1/3 & \text{the radiation field or relativistic particles} \\ 0 & \text{the nonrelativistic matter or particles} \\ -1 & \text{the cosmological constant}\,\Lambda \end{cases} \tag{12.1.40}$$

Note that the density ρ in Eqs. (12.1.14), (12.1.15) and (12.1.17) includes neither the vacuum energy density as given by the cosmological constant Λ nor the curvature term. In the case that a specific energy density ρ_j dominates all else, one substitutes Eq. (12.1.39) into Eq. (12.1.17) and then obtains

$$\dot{\rho}_j = -3H(1+w_j)\rho_j = -3(1+w_j)\rho_j \frac{\dot{R}}{R} \tag{12.1.41}$$

As a result,

$$\rho_j \propto R^{-3(1+w_j)} \tag{12.1.42}$$

Substituting Eq. (12.1.42) into the Friedmann equation in Eq. (12.1.14) and neglecting the Λ and curvature terms, we find

$$R(t) \propto t^{\frac{2}{3(1+w_j)}} \tag{12.1.43}$$

Obviously, this result is only valid during a specific period in the evolution of the Universe (i.e. during which the curvature and vacuum energy Λ terms can be negligible).

12.1.5.2. *The solutions of the standard model of the Universe for the epoch dominated by radiation*

The early Universe is in a state of high temperature and high density, and thus all the particles were relativistic. In other words, this epoch is dominated by the radiation field and relativistic particles, corresponding to $w_{\rm r} = 1/3$ in Eq. (12.1.40). Then Eq. (12.1.42) leads us to

$$\rho_{\rm r} \propto R^{-4} \tag{12.1.44}$$

Now that the curvature term is proportional to R^{-2} and the Λ term is also tiny, so they are both negligible as compared with $\rho_{\rm r} \propto R^{-4}$. Given the fact that the volume of the Universe is proportional to $\propto R^3$, one might naively expect that $\rho_{\rm r} \propto R^{-3}$ holds. This is not true, because the expansion of the Universe gives rise to the redshift effect, which not only makes the wavelength λ become larger and the frequency ν become smaller but also results in $E_{\rm r} \propto \nu \propto R^{-1}$. This point can also be seen from Eq. (12.1.34). That is why $\rho_{\rm r} \propto R^{-4}$ holds. Substituting $w_{\rm r} = 1/3$ into Eq. (12.1.43), we are left with

$$R(t) \propto t^{1/2}, \quad H = \frac{1}{2t} \tag{12.1.45}$$

12.1.5.3. *The solutions of the standard model of the Universe for the epoch dominated by matter*

As the Universe expands and its temperature decreases, the density of non-relativistic matter will eventually be dominant over the density of radiation. During this epoch the pressure of matter is vanishing, as shown in Eq. (12.1.40). Therefore,

$$\rho_{\rm m} \propto R^{-3} \tag{12.1.46}$$

Since the κ and Λ terms are both negligible in this era, one may solve Eq. (12.1.43) and then obtain

$$R(t) \propto t^{2/3}, \quad H = \frac{2}{3t} \tag{12.1.47}$$

12.1.5.4. *The solutions of the standard model of the Universe for the epoch dominated by vacuum energy*

The cosmological constant Λ is actually a measure of the vacuum energy. Let us define

$$\begin{cases} \Lambda = 8\pi G_{\mathrm{N}}\rho_{\mathrm{V}} \\ \rho_{\mathrm{V}} = \dfrac{\Lambda}{8\pi G_{\mathrm{N}}} \end{cases} \tag{12.1.48}$$

When the curvature term is neglected, the Friedmann equation can be rewritten as

$$\left(\frac{\dot{R}}{R}\right)^2 = \frac{\Lambda}{3} \tag{12.1.49}$$

from which we find

$$R(t) = \mathrm{e}^{\sqrt{\frac{\Lambda}{3}}t}, \quad H = \sqrt{\frac{\Lambda}{3}} \tag{12.1.50}$$

During this epoch the Universe expands exponentially. Here the keypoint is that the vacuum energy Λ provides a negative pressure:

$$p = w\rho_\Lambda = -\frac{\Lambda}{8\pi G_{\mathrm{N}}} \tag{12.1.51}$$

Of course, the value of w in the equation of state is unnecessarily equal to a constant (-1). A more general case is that the vacuum energy undergoes a sort of dynamic evolution. This kind of vacuum energy (i.e. the so-called dark energy) will be discussed later.

12.1.6. ***The age of the Universe***

According to the standard ΛCDM model, the age of the Universe is finite. Equation (12.1.18) leads to

$$\frac{8\pi G_{\mathrm{N}}}{3} = \frac{H^2(t)}{\rho_{\mathrm{c}}(t)} = \frac{H^2(t_0)}{\rho_{\mathrm{c}}(t_0)} \tag{12.1.52}$$

where t_0 denotes the present time. Note that the quantity on the left-hand side of Eq. (12.1.52) is time-independent. Given Eq. (12.1.14),

$$\begin{aligned} H^2(t) &= \frac{8\pi G_{\mathrm{N}}}{3}\rho - \frac{\kappa}{R^2} + \frac{\Lambda}{3} \\ &= \frac{8\pi G_{\mathrm{N}}}{3}(\rho_{\mathrm{m}} + \rho_{\mathrm{r}} + \rho_\Lambda) - \frac{\kappa}{R^2} \end{aligned} \tag{12.1.53}$$

With the help of Eq. (12.1.52), we obtain

$$H^2(t) = \frac{H^2(t_0)}{\rho_{\rm c}(t_0)}[\rho_{\rm m}(t) + \rho_{\rm r}(t) + \rho_\Lambda(t) - \frac{\rho_{\rm c}(t_0)}{H^2(t_0)}\frac{\kappa}{R^2(t)}] \tag{12.1.54}$$

Equation (12.1.37) tells us that

$$R(t) = \frac{R(t_0)}{1+z} \tag{12.1.55}$$

Taking account of $\rho_{\rm m}(t) \propto R(t)^{-3}$, $\rho_{\rm r}(t) \propto R(t)^{-4}$ and $\rho_\Lambda(t) = \text{constant}$ given in Sec. 12.1.5, we rewrite Eq. (12.1.54) as

$$\begin{aligned} H^2(t) = \frac{H^2(t_0)}{\rho_{\rm c}(t_0)}[&\rho_{\rm m}(t_0)(1+z)^3 + \rho_{\rm r}(t_0)(1+z)^4 \\ &+\rho_\Lambda(t_0) - \frac{\rho_{\rm c}(t_0)}{H^2(t_0)}\frac{\kappa}{R^2(t_0)}(1+z)^2] \end{aligned} \tag{12.1.56}$$

With the help of Eq. (12.1.25), we substitute

$$\frac{\kappa}{H^2(t_0)R^2(t_0)} = \Omega_{\rm tot}(t_0) - 1$$

into Eq. (12.1.56) and arrive at

$$\begin{aligned} H^2(t) = H^2(t_0)[&\Omega_{\rm m}(t_0)(1+z)^3 + \Omega_{\rm r}(t_0)(1+z)^4 + \Omega_\Lambda(t_0) \\ &+(1 - \Omega_{\rm m}(t_0) - \Omega_{\rm r}(t_0) - \Omega_\Lambda(t_0))(1+z)^2] \end{aligned} \tag{12.1.57}$$

Using Eq. (12.1.55) and keeping in mind that $R(t_0)$ is a constant, we find

$$H(t) = \frac{\dot{R}(t)}{R(t)} = -\frac{1}{1+z}\frac{{\rm d}z}{{\rm d}t} \tag{12.1.58}$$

and

$${\rm d}t = -\frac{{\rm d}z}{(1+z)H(t)} \tag{12.1.59}$$

Given $z = 0$ at $t = t_0$ and $z = \infty$ at $t = 0$, the age t_0 of the Universe can be expressed as

$$t_0 = \int_0^\infty \frac{{\rm d}z}{(1+z)H(t)} \tag{12.1.60}$$

Since the space of our Universe is almost flat, it is reasonable to take $\kappa = 0$ and obtain $\Omega_{\rm tot} = \Omega_{\rm m} + \Omega_{\rm r} + \Omega_{\rm V} \approx 1$. Note that $\Omega_{\rm r}(t_0)$ is negligibly small, as one can see from Eq. (12.1.26). So Eq. (12.1.57) can be simplified to

$$H^2(t) \approx H^2(t_0)[\Omega_{\rm m}(t_0)(1+z)^3 + \Omega_\Lambda(t_0)] \tag{12.1.61}$$

Substituting this result into Eq. (12.1.60), we have

$$\begin{aligned} t_0 &= \frac{1}{H(t_0)} \int_0^\infty {\rm d}z \frac{1}{(1+z)\sqrt{\Omega_{\rm m}(t_0)(1+z)^3 + \Omega_\Lambda(t_0)}} \\ &= \frac{1}{3H(t_0)\sqrt{\Omega_\Lambda(t_0)}} \ln \frac{1+\sqrt{\Omega_\Lambda(t_0)}}{1-\sqrt{\Omega_\Lambda(t_0)}} \end{aligned} \tag{12.1.62}$$

where the approximate relation $\Omega_\Lambda(t_0) \approx 1 - \Omega_{\rm m}(t_0)$ has been used. Taking account of the experimental values $\Omega_{\rm m}(t_0) \approx 0.315$ and $\Omega_\Lambda(t_0) \approx 0.685$ in Eq. (12.1.26), we get

$$t_0 \approx \frac{0.956}{H(t_0)} \sim \frac{1}{H(t_0)} \tag{12.1.63}$$

Substituting $H(t_0) = 100h_0\,{\rm km~s}^{-1}M_{\rm pc}^{-1}$ into the above equation, where $h_0 = 0.673 \pm 0.012$ and $M_{\rm pc} = 10^6 \times 3.09 \times 10^{13}{\rm km}$, we are left with $t_0 \approx 13.9\,{\rm Gyr}$, a result compatible with the value $t_0 = 13.81 \pm 0.05\,{\rm Gyr}$ measured by the Planck Collaboration in 2013 [12].

12.2. Hot Big Bang and Radiation in the Early Universe

The radius of the very early Universe is very small, but its temperature is extremely high. In Sec. 12.1.5 we have shown $\rho_{\rm r} \propto R^{-4}(t)$ for the radiation field and $\rho_{\rm m} \propto R^{-3}(t)$ for the matter field. So at this stage the $\rho_{\rm m}$, κ and Λ terms in the Friedmann equation are all negligible, leading us to

$$\frac{\dot{R}^2}{R^2} = \frac{8\pi G_{\rm N}}{3}\rho_{\rm r} \tag{12.2.1}$$

Given $\rho_{\rm r} \propto R^{-4}$, we have

$$\frac{\dot{\rho}_{\rm r}}{\rho_{\rm r}} = -\frac{4\dot{R}}{R} = -4\left(\frac{8\pi G_{\rm N}}{3}\rho_{\rm r}\right)^{1/2} \tag{12.2.2}$$

As a result,

$$\rho_{\mathrm{r}} = \frac{3}{32\pi G_{\mathrm{N}}} \frac{1}{t^2} \tag{12.2.3}$$

First of all, let us consider the era of radiation during which the photons play the dominant role and the contributions from other relativistic particles are neglected. In this case, the gas of photons is in equilibrium and their energy density is given by

$$\rho_\gamma = 4\sigma_{\mathrm{SB}} T^4 = \frac{g_\gamma \pi^2}{30} T^4 \tag{12.2.4}$$

where $\sigma_{\mathrm{SB}} = \pi^2/60$ is the Stefan–Boltzmann constant and $g_\gamma = 2$ is the degrees of freedom for the spin of a photon. One may also see

$$\rho_\gamma \propto R^{-4} \to T \propto R^{-1}$$

A combination of Eqs. (12.2.3) and (12.2.4) leads to

$$T = \left(\frac{45}{16 g_\gamma \pi^3 G_{\mathrm{N}}}\right)^{1/4} \frac{1}{\sqrt{t}} \tag{12.2.5}$$

Here we have adopted the natural units (i.e. $k = \hbar = c = 1$ with k being the Boltzmann constant) by which $G_{\mathrm{N}} = 6.70881 \times 10^{-45}\ \mathrm{MeV}^{-2}$ holds. Substituting this result into Eq. (12.2.5) and taking account of $\hbar^2 = (6.58 \times 10^{-22}\,\mathrm{MeV\ s})^2$, we obtain

$$T \approx 1.31\ \mathrm{MeV}/\sqrt{t} \tag{12.2.6}$$

where t is in unit of second. Inputting $k = 8.62 \times 10^{-11}\ \mathrm{MeV\ K}^{-1}$, we arrive at

$$T \approx 1.52 \times 10^{10}\,\mathrm{K}/\sqrt{t} \tag{12.2.7}$$

One can see $T \to \infty$ for $t \to 0$, indicating that the Big Bang of the Universe should begin with an extremely high temperature. That is why, the Big Bang model is also referred to as the “hot Big Bang model”.

With the help of Eqs. (12.2.6) and (12.2.7), one may estimate today’s temperature of the CMB (Cosmic Microwave Background) radiation.

Taking $t_0 \sim 13.81\,\mathrm{Gyr} \sim 4.36 \times 10^{17}\,\mathrm{s}$, one is left with

$$T \approx 23\,\mathrm{K} \tag{12.2.8}$$

The COBE satellite has done very precise measurements of the CMB temperature and reported the result [13]

$$T = 2.7255 \pm 0.0006\,\mathrm{K} \tag{12.2.9}$$

The first measurement of this kind was done by Penzias and Wilson in 1965 [14]. Figure 12.1 shows the observed CMB spectrum, which is in good agreement with the spectrum expected from the black-body radiation at temperature $T = 2.726\,\mathrm{K}$.

It is worth pointing out that the estimate of the CMB temperature made in Eq. (12.2.8) is far above its experimental value. Given $\rho_\gamma \propto T^4 \propto R^{-4}$, the temperature $T \propto R^{-1}$ should decrease in proportion to $t^{-1/2}$. During the matter-dominated epoch of the Universe, the temperature T decreases much faster, in proportion to $t^{-2/3}$ as one can see from Eq. (12.1.47). After these issues are taken into account, a careful estimate of the CMB temperature will lead us to a result lower than that given in Eq. (12.2.8).

Equation (12.2.4) allows us to calculate the number density of photons N_γ for the CMB radiation at temperature $T = 2.726\,\mathrm{K}$. The energy of

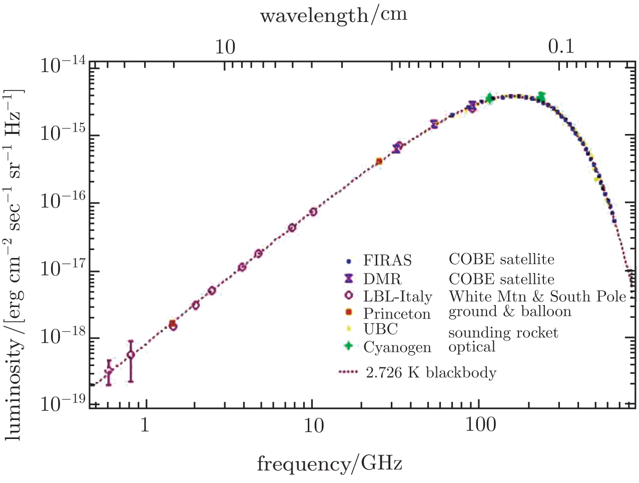


Fig. 12.1. The observed spectrum of CMB radiation at T=2.726 K. The vertical and horizontal axes stand for the energy density multiplied by the speed of light and the frequency, respectively.

photons from the black-body radiation is given by $E = pc = h\nu$, and their number density satisfies the Bose–Einstein distribution

$$N(E)\mathrm{d}p = \frac{p^2 \mathrm{d}p}{\pi^2[\exp(E/T) - 1]}\left(\frac{g_\gamma}{2}\right) \tag{12.2.10}$$

Integrating the above equation over energy E, one obtains[a]

$$N_\gamma = \frac{2\zeta(3)}{\pi^2}\frac{g_\gamma}{2}T^3 \approx 411\left(\frac{T}{2.726\mathrm{K}}\right)^3 \mathrm{cm}^{-3} \tag{12.2.11}$$

The number density of photons at the CMB temperature (i.e. $T = 2.726$K) turns out to be

$$N_\gamma \approx 411\,\mathrm{cm}^{-3} \tag{12.2.12}$$

At the same temperature, one may also figure out the energy density of those CMB photons from Eq. (12.2.4):

$$\rho_\gamma = \frac{\pi^2}{15}\frac{(kT)^4}{\hbar^3 c^3} = 0.261\,\mathrm{MeV\,m}^{-3}$$

or

$$\rho_\gamma/c^2 = 4.65 \times 10^{-31}\,\mathrm{kg\,m}^{-3} \tag{12.2.13}$$

The value of Ω_γ at the present time t_0 is therefore given by

$$\Omega_\gamma(t_0) = \rho_\gamma(t_0)/\rho_\mathrm{c} \approx 4.84 \times 10^{-5} \tag{12.2.14}$$

which is far below the energy density of matter today, $\Omega_\mathrm{m}(t_0) \approx 0.315$.

Note that Eq. (12.2.4) only contains the contribution of photons to the energy density of the Universe. In general, any relativistic particles may contribute to ρ_γ if and only if they are stable enough.

---

[a] The following integrals are closely associated with the Bose–Einstein and Fermi–Dirac distributions:

$$\int_0^\infty \frac{x^2\mathrm{d}x}{\mathrm{e}^x - 1} = 2\zeta(3) = 2.404,$$

$$\int_0^\infty \frac{x^3\mathrm{d}x}{\mathrm{e}^x - 1} = \frac{\pi^4}{15}, \int_0^\infty \frac{x^3\mathrm{d}x}{\mathrm{e}^x + 1} = \frac{7}{8}\frac{\pi^4}{15},$$

$$\int_0^\infty \frac{x^2\mathrm{d}x}{\mathrm{e}^x + 1} = \frac{3}{2}\zeta(3) = \frac{3}{4} \times 2.404,$$

where $\zeta(x)$ is the Riemann zeta function.

For relativistic bosons, one may simply replace g_γ in Eq. (12.2.4) with $g_{\rm b}$, where $g_{\rm b}$ denotes the number of spin states of a boson. Of course, $g_\gamma = 2$ for a photon.

For relativistic fermions, one may approximately describe them with the fermion gas, which satisfies the Fermi–Dirac distribution. The number of fermions per unit volume can be calculated by means of

$$N(E){\rm d}p = \frac{p^2{\rm d}p}{\pi^2[\exp(E/T)+1]}\left(\frac{g_{\rm f}}{2}\right) \tag{12.2.15}$$

because $E = \sqrt{p^2+m^2} \approx \sqrt{p^2} = p$. The corresponding number density of fermions is given by

$$n_{\rm f} = \int_0^\infty \left(\frac{g_{\rm f}}{2}\right)\frac{p^2{\rm d}p}{\pi^2[\exp(E/T)+1]} = \frac{3\zeta(3)}{2\pi^2}\left(\frac{g_{\rm f}}{2}\right)T^3 \tag{12.2.16}$$

and the energy density is obtained as

$$\begin{aligned}\rho_{\rm f} &= \int_0^\infty EN(E){\rm d}p \approx \int_0^\infty \frac{p^3{\rm d}p}{\pi^2[\exp(E/T)+1]}\left(\frac{g_{\rm f}}{2}\right)\\ &= \frac{7\pi^2}{240}g_{\rm f}T^4 = \frac{7}{8}\frac{\pi^2}{30}g_{\rm f}T^4 \end{aligned}\tag{12.2.17}$$

Taking account of the contribution of those reasonably stable relativistic particles to the energy density of radiation, we have

$$\begin{aligned}\rho_{\rm r} = \rho_{\rm b} + \rho_{\rm f} &= \sum_{\rm b}\int_0^\infty \frac{g_{\rm b}Ep^2{\rm d}p}{2\pi^2[\exp(E/T)-1]} + \sum_{\rm f}\int_0^\infty \frac{g_{\rm f}Ep^2{\rm d}p}{2\pi^2[\exp(E/T)+1]}\\ &= \frac{\pi^2}{30}g_* T^4\end{aligned}\tag{12.2.18}$$

where

$$g_* = \sum_{\rm b} g_{\rm b} + \frac{7}{8}\sum_{\rm f} g_{\rm f} \tag{12.2.19}$$

Given extremely high temperatures in the very early Universe, all the fundamental particles described by the standard ${\rm SU(3)_C \times SU(2) \times U(1)}$ gauge theory are relativistic. So they contribute to $\rho_{\rm r}$ in Eq. (12.2.18); namely, g_* should be computed according to Eq. (12.2.19). In this case g_γ in Eqs. (12.2.4) and (12.2.5) should be replaced by g_*; namely, Eq. (12.2.4) should be replaced by Eq. (12.2.18), and Eq. (12.2.5) should be modified as

$$T = \left(\frac{45}{16g_* \pi^3 G_{\rm N}}\right)^{1/4} \frac{1}{\sqrt{t}} \tag{12.2.20}$$

from which we obtain

$$t = \left(\frac{45}{16g_* \pi^3}\right)^{1/2} \frac{M_{\rm Pl}}{T^2} \approx \frac{0.301}{\sqrt{g_*}} \frac{M_{\rm Pl}}{T^2} \tag{12.2.21}$$

where $M_{\rm Pl} = 1/\sqrt{G_{\rm N}} \approx 1.22 \times 10^{19}$GeV is the Planck mass. Now let us look at what value g_* should take. We first consider the era in which the temperature of the Universe is so high that the electroweak phase transition has not taken place. The massless electroweak gauge fields include W^1, W^2, W^3, B, eight gluons and the Higgs doublet with two complex components. A sum of their spin states is equal to 28 (i.e. $\sum_{\rm b} g_{\rm b} = 28$). In the quark sector, the number of spin states for six quarks and six antiquarks is 72, because each quark (or antiquark) has three different colors. In the lepton sector, the number of spin states for three charged leptons and their antiparticles is 12, and that for three massless neutrinos and their antiparticles is 6. In total $\sum_{\rm f} g_{\rm f} = 90$ in the fermion sector. We are therefore left with $g_* = 28 + \frac{7}{8} \times 90 = 106.75$ during this epoch of the Universe. The electroweak phase transition takes place at $T \sim 200$ GeV, where the Higgs boson and some of the gauge bosons acquire their masses, but the photon and gluons remain massless. Note that the degrees of freedom for γ, $\mathrm{W}^{\pm}$, Z and H are the same as those for W^1, W^2, W^3, B and ϕ. So $\sum_{\rm b} g_{\rm b} = 28$ and $\sum_{\rm f} g_{\rm f}$ keep unchanged, and hence $g_* = 106.75$ holds. When the temperature goes down with the expansion of the Universe and becomes lower than $m_{\rm W}$ and $m_{\rm Z}$ (or equivalently, kT is smaller than $m_{\rm W}$ and $m_{\rm Z}$), both W and Z will decay quickly because their lifetimes are smaller than 10^{-23}s. In this case W, Z and H cannot contribute to g_* anymore. When $T < \Lambda_{\rm QCD} \sim 200\,{\rm MeV}$, quarks, antiquarks and quark–gluon plasma do not exist anymore. Then one is left with only pions, nucleons and other hadrons, but only protons and neutrons can survive because other hadrons all decay within a very short time ($10^{-9} - 10^{-8}$s). During this epoch the rates for the bidirectional reactions $\gamma + \gamma \leftrightarrows \mathrm{N} + \bar{\mathrm{N}}$ are not equal anymore. The probability for $\mathrm{N}\bar{\mathrm{N}}$ to annihilate into photons is much larger than the probability of $\gamma + \gamma \to \mathrm{N} + \bar{\mathrm{N}}$, making most of $\mathrm{N}\bar{\mathrm{N}}$ disappear. When $T < 20\,{\rm MeV}$, only a small fraction of nucleons can still survive.

Now let us use Eq. (12.2.21) to estimate the time corresponding to each of the energy scales discussed above. Rewriting Eq. (12.2.21) as

$$tT^2 \approx \frac{0.301}{\sqrt{g_*}} M_{\rm Pl} \tag{12.2.22}$$

and inputting the value of $M_{\rm Pl}$, one arrives at

$$tT^2 \approx \frac{2.42}{\sqrt{g_*}} {\rm MeV}^2 \cdot {\rm s} \tag{12.2.23}$$

Taking $g_* = 106.75$, one may find that the Planck scale $T \sim 10^{19}\,{\rm GeV}$ corresponds to $t \sim 10^{-45}{\rm s}$ after the Big Bang. Similarly, the GUT scale $T \sim 10^{16}\,{\rm GeV}$ and the Fermi scale $T \sim 10^{2}{\rm GeV}$ correspond, respectively, to $t \sim 10^{-39}{\rm s}$ and $t \sim 10^{-11}{\rm s}$ after the Big Bang. When $T \sim \Lambda_{\rm QCD} \sim 200\,{\rm MeV}$, one cannot take the value of g_* to be 106.75 anymore. At this energy scale we may assume that the photon, eight gluons, three light quarks (u,d,s), e, μ, $\nu_{\rm e}$, ν_μ, ν_τ and their antiparticles contribute to the value of g_*. As a result, $g_* = 18 + \frac{7}{8} \times 50 = 61.75$. Then we find that $T \sim \Lambda_{\rm QCD}$ corresponds to $t \sim 7.7 \times 10^{-6}{\rm s}$.

We refer the reader to Chapter 9 of Ref. [16] and Chapter 10 of Ref. [17] for similar discussions.

12.3. Neutrino Decoupling and Cosmic Neutrino Background

Now let us proceed to discuss when neutrinos and antineutrinos decouple from the cosmic thermal bath and how such relic neutrinos form the cosmic neutrino background (CνB) in today's Universe [16].

When the temperature of the Universe T decreases to a few MeV, the charged leptons μ and τ have also decoupled from the thermal bath. At this time, the only survival particles are γ, ${\rm e}^\pm$, ν_α and $\bar{\nu}_\alpha$ (for $\alpha = {\rm e}, \mu, \tau$). The number densities of these particles are comparable in thermal equilibrium, in which the bidirectional reactions $\gamma + \gamma \leftrightarrows {\rm e}^+ + {\rm e}^- \leftrightarrows \nu_\alpha + \bar{\nu}_\alpha$ (for $\alpha = {\rm e}, \mu, \tau$) take place. But the reaction ${\rm e}^+ + {\rm e}^- \leftrightarrows \nu_\alpha + \bar{\nu}_\alpha$ can only occur via the weak neutral–current interactions, and its average cross section is given by $\langle \sigma_{{\rm e}^\pm} v_{{\rm e}^\pm} \rangle \sim G_{\rm F}^2 T^2$, where the notation $\langle \ \rangle$ means an average taken with respect to the velocities of electrons and positrons. The probability of this bidirectional reaction is $W = n_{{\rm e}^\pm} \langle \sigma_{{\rm e}^\pm} v_{{\rm e}^\pm} \rangle$, where $n_{{\rm e}^\pm}$ denotes the number density of ${\rm e}^\pm$. In view of Eq. (12.2.16), one obtains $n_{{\rm e}^\pm} \propto T^3$ and $W \approx G_F^2 T^5$. During the epoch dominated by radiation, the Hubble expansion rate of the Universe is given by $H = \frac{\dot{R}}{R} = \frac{1}{2t}$ in Eq. (12.1.45). Taking account of Eq. (12.2.21), one has

$$H(t) = \frac{1}{2t} = 1.66\sqrt{g_*}\frac{T^2}{M_{\mathrm{Pl}}} \tag{12.3.1}$$

When the reaction rate of $\mathrm{e}^+ + \mathrm{e}^- \leftrightarrows \nu_\alpha + \bar{\nu}_\alpha$ becomes smaller than the Hubble expansion rate (i.e. $W < H$), or equivalently $G_{\mathrm{F}}^2 T^5 < 1.66\frac{\sqrt{g_*}T^2}{M_{\mathrm{Pl}}}$, neutrinos and antineutrinos are decoupled from the thermal bath. At this moment, the corresponding temperature reads

$$T < \left(\frac{\sqrt{g_*}}{G_{\mathrm{F}}^2 M_{\mathrm{Pl}}}\right)^{1/3} = 1.25\ \mathrm{MeV} \approx 1.5 \times 10^{10}\ \mathrm{K} \tag{12.3.2}$$

where $g_* = 10.75$ includes the contributions from γ, $\mathrm{e}^\pm$, ν_α and $\bar{\nu}_\alpha$. Equation (12.3.2) leads us to the approximate neutrino-decoupling temperature

$$T_{\mathrm{fr}} \approx 1\mathrm{MeV} \tag{12.3.3}$$

When $T < T_{\mathrm{fr}}$ holds, neutrinos and antineutrinos travel freely in the space of the Universe. The temperature (or frequency) of such massless relativistic particles and antiparticles decreases with the expansion of the Universe in a way of $T_\nu \propto 1/R$.

After neutrinos and antineutrinos are decoupled from the thermal bath, the temperature of the Universe continues to go down. When $T \leq m_{\mathrm{e}}$ holds, neutrinos, $\mathrm{e}^\pm$ and γ form a mixed gas in the Universe and share the same temperature. At this moment, it is reasonable to assume that the reaction rate of $\mathrm{e}^+\mathrm{e}^- \to \gamma\gamma$ begins to exceed far from that of $\gamma\gamma \to \mathrm{e}^+\mathrm{e}^-$, leading us to $T_\nu = T$ and $g_* = 2 + \frac{7}{8} \times 4 = \frac{11}{2}$. The entropy density of the Universe can be calculated by using $S = \int \frac{\mathrm{d}Q}{T}$, where Q denotes the energy density of the mixed gas of photons, electrons and positrons. Equation (12.2.18) tells us that $Q = \frac{\pi^2 g_*}{30}T^4$ and $g_* = 11/2$, and thus

$$S = \int \frac{\mathrm{d}Q}{T} = \int \frac{2g_*\pi^2}{15}T^2\mathrm{d}T = \frac{2g_*\pi^2}{45}T^3 = \frac{11}{4}\frac{2g_\gamma\pi^2}{45}T^3 \tag{12.3.4}$$

can be obtained.

Because of $\mathrm{e}^+\mathrm{e}^- \to \gamma\gamma$, the number of $\mathrm{e}^\pm$ goes down quickly. In this case, one may assume that the instantaneous entropy density is mainly ascribed to the contribution of photons, whose temperature T_γ is somewhat enhanced thanks to the reaction $e^+e^- \to \gamma\gamma$. The entropy density of photons turns out to be

$$S_\gamma = \int \frac{2g_\gamma \pi^2}{15} T_\gamma^2 \mathrm{d}T_\gamma = \frac{2g_\gamma \pi^2}{45} T_\gamma^3 \tag{12.3.5}$$

where $Q = \rho_\gamma$ has been taken. Since the above reaction has been assumed to happen instantaneously and the instantaneous expansion of the Universe has been neglected, the entropy conservation of the Universe as an isolated system assures that its entropy density keeps unchanged. As a result,

$$S_\gamma = S \tag{12.3.6}$$

With the help of Eqs. (12.3.4) and (12.3.5), we find

$$T_\gamma = \left(\frac{11}{4}\right)^{1/3} T = \left(\frac{11}{4}\right)^{1/3} T_\nu \tag{12.3.7}$$

and

$$T_\nu \equiv T = \left(\frac{4}{11}\right)^{1/3} T_\gamma \tag{12.3.8}$$

Equation (12.3.7) tells us that the temperature of photons has really been enhanced. In the subsequent era of the Universe with $T \ll 1\,\mathrm{MeV}$, the total energy density of relativistic particles is

$$\begin{aligned} \rho_{\gamma+\nu} = \rho_\gamma + \rho_\nu &= \frac{g_\gamma \pi^2}{30} T_\gamma^4 + \frac{7}{8}\frac{g_\nu \pi^2}{30} T_\nu^4 \\ &= \frac{g'_* \pi^2}{30} T_\gamma^4 \end{aligned} \tag{12.3.9}$$

where

$$g'_* = g_\gamma + \frac{7}{8} g_\nu \left(\frac{T_\nu}{T_\gamma}\right)^4 = 2 + 6 \times \frac{7}{8} \left(\frac{4}{11}\right)^{4/3} \approx 3.36 \tag{12.3.10}$$

Later on the relic neutrinos have nothing to do with the photons, but they undergo the same redshift effect caused by the expansion of the Universe. That is why, the ratio of their number densities should keep invariant.

In other words, the ratio n_ν/n_γ keeps unchanged from the neutrino decoupling until today. Taking account of Eqs. (12.2.16) and (12.2.11), we obtain

$$\begin{aligned} n_\nu &= \frac{3\zeta(3)}{2\pi^2}\left(\frac{g_\nu}{2}\right)T_\nu^3 = \frac{3}{4}\left(\frac{g_\nu}{g_\gamma}\right)\left(\frac{T_\nu}{T_\gamma}\right)^3 n_\gamma \\ &= \frac{9}{11}n_\gamma \approx 336\left(\frac{T_\gamma}{2.726\mathrm{K}}\right)^3 \mathrm{cm}^{-3} \end{aligned} \tag{12.3.11}$$

where $g_\nu = 6$ has been input. So today's CMB temperature is $T_\gamma \sim 2.726\,\mathrm{K}$, the corresponding C$\nu$B temperature is $T_\nu \sim 1.946\,\mathrm{K}$, and the number density of relic neutrinos is $n_\nu \sim 336\,\mathrm{cm}^{-3}$. How to measure such a CνB of the Universe is a great challenge in astronomy and cosmology. In this connection, we refer the reader to Ref. [16] for some further discussions.

12.4. Big Bang Nucleosynthesis and Abundances of Primordial Light Elements

The Big Bang nucleosynthesis (BBN) takes place when the temperature T of the Universe lies in the range of 1–0.1 MeV. In this epoch which is dominated by radiation, Eq. (12.2.23) remains valid but $g_* \sim 10.75$ (when $T \sim 1\,\mathrm{MeV}$, the contributions include γ, e, ν_α and $\bar{\nu}_\alpha$) or $g'_* \sim 3.36$ (when $T \sim 0.1\,\mathrm{MeV}$, the contributions include γ, ν_α and $\bar{\nu}_\alpha$) should be carefully taken into account. The corresponding time interval is $t \sim 0.74\,\mathrm{s}$ ($T \sim 1\,\mathrm{MeV}$) — $132\,\mathrm{s}$ ($T \sim 0.1\,\mathrm{MeV}$), which can be calculated from Eq. (12.2.23). In Sec. 12.2 we have pointed out that only a small fraction of protons and neutrons can still survive and stay in thermal equilibrium when the temperature of the Universe becomes lower than 20 MeV. When $T \leq 1\,\mathrm{MeV}$, the electron (anti)neutrinos make protons and neutrons keep in thermal equilibrium via the charged weak interaction reactions

$$\mathrm{n} + \nu_\mathrm{e} \Leftrightarrow \mathrm{p} + \mathrm{e}^-, \quad \mathrm{p} + \bar{\nu}_\mathrm{e} \Leftrightarrow \mathrm{n} + \mathrm{e}^+, \quad \mathrm{n} \Leftrightarrow \mathrm{p} + \mathrm{e}^- + \bar{\nu}_\mathrm{e} \tag{12.4.1}$$

before they themselves are decoupled from the thermal bath. In this case, the numbers of protons and neutrons should be proportional to the Boltzmann factor $\mathrm{e}^{-m/T}$. Therefore,

$$\frac{N_\mathrm{n}}{N_\mathrm{p}} \sim \mathrm{e}^{-(M_\mathrm{n}-M_\mathrm{p})/T} = \mathrm{e}^{-Q/T}, \quad Q = M_\mathrm{n} - M_\mathrm{p} \tag{12.4.2}$$

Here we have ignored the velocities of nucleons by assuming them to be almost at rest. This assumption is reasonable for $T{\sim}1 \sim 0.1$ MeV. A more careful calculation yields the freeze-out temperature of neutrino decoupling $T_{\mathrm{fr}}{\sim}0.72$ MeV [18]. Substituting this value into Eq. (12.4.2), we obtain

$$\frac{N_{\mathrm{n}}}{N_{\mathrm{p}}}{\sim}\mathrm{e}^{-1.293/0.72} \approx 1.66 \sim 1/6 \tag{12.4.3}$$

and the corresponding time from the Big Bang is $t_{\mathrm{fr}} \approx 1.4$s. When the temperature of the Universe is below $T_{\mathrm{fr}} \sim 0.72\,\mathrm{MeV}$, the number of neutrons will decrease via the β decay $\mathrm{n} \to \mathrm{p} + \mathrm{e}^- + \bar{\nu}_{\mathrm{e}}$ until they are bound in deuterium to start the nucleosynthesis process. At this moment, the production of neutrons via Eq. (12.4.1) is essentially stopped. That is why for the duration t' after the neutrino decoupling (i.e. the time since $T_{\mathrm{fr}} \approx 1.4$s) the neutron-to-proton ratio is given as follows (note that the numbers of neutron and protons are $N_{\mathrm{n}}(0)\mathrm{e}^{-t'/\tau_{\mathrm{n}}}$ and $N_{\mathrm{p}}(0)+N_{\mathrm{n}}(0)(1-\mathrm{e}^{-t'/\tau_{\mathrm{n}}})$, respectively):

$$\frac{N_{\mathrm{n}}}{N_{\mathrm{p}}} \sim \frac{\mathrm{e}^{-Q/T_{\mathrm{fr}}}\mathrm{e}^{-t'/\tau_{\mathrm{n}}}}{1 + \mathrm{e}^{-Q/T_{\mathrm{fr}}}\left(1 - \mathrm{e}^{-t'/\tau_{\mathrm{n}}}\right)} \tag{12.4.4}$$

where $\tau_{\mathrm{n}} = (880.3 \pm 1.1)$s is the lifetime of neutrons.

Using Eq. (12.2.23) and taking $t{\sim}160$ s (i.e. $t' = 160 - 1.4 = 158.6$s) in accordance with $g_* = g'_* = 3.36$, we find that the corresponding temperature is $T{\sim}0.09$ MeV. At this time the nucleosynthesis process almost stopped. In view of Eq. (12.4.4), we get

$$\frac{N_{\mathrm{n}}}{N_{\mathrm{p}}}{\sim}1.4 \tag{12.4.5}$$

In fact, the nucleosynthesis has started since $T{\sim}1$ MeV, and it is mainly realized via the synthesis of deuterium. Namely,

$$\mathrm{n} + \mathrm{p} \Leftrightarrow \mathrm{D} + \gamma + E_{\mathrm{bind}} \tag{12.4.6}$$

where $E_{\mathrm{bind}} = 2.22$ MeV is the binding energy of deuterium D.

The reaction in Eq. (12.4.6) is bidirectional. Because it takes place via electromagnetic interactions, its cross section is relatively large (about 0.1 mb). That is why this process can stay in thermal equilibrium. Given a huge amount of photons existing in the Universe, the deuterium photodissociation will not stop until $T{\sim}0.09$ MeV. But at this moment the synthesis of helium takes place. That is,

$$\begin{cases} \mathrm{D} + \mathrm{n} \to {}^3\mathrm{H} + \gamma, \mathrm{D} + \mathrm{D} \to {}^3\mathrm{H} + \mathrm{n} \\ {}^3\mathrm{H} + \mathrm{p} \to {}^4\mathrm{He} + \gamma \\ \mathrm{D} + \mathrm{p} \to {}^3\mathrm{He} + \gamma, \mathrm{D} + {}^3\mathrm{H} \to {}^4\mathrm{He} + \mathrm{n} \\ {}^3\mathrm{He} + \mathrm{n} \to {}^4\mathrm{He} + \gamma, \mathrm{D} + {}^3\mathrm{He} \to {}^4\mathrm{He} + \mathrm{p} \end{cases} \tag{12.4.7}$$

Neutrons do not decay anymore after they are bounded in ^{4}He, making the ratio $N_\mathrm{n}/N_\mathrm{p}$ keep unchanged. The corresponding time is about 160 s after the Big Bang, as one can estimate by using Eq. (12.2.23). If all the neutrons are bound in stable ^{4}He and do not decay anymore, one may calculate the ratio of the primordial ^{4}He mass to the total mass of all the elements:

$$Y_\mathrm{p} = \frac{M_\mathrm{He} N_\mathrm{He}}{M_\mathrm{N}(N_\mathrm{n} + N_\mathrm{p})} \tag{12.4.8}$$

Note that the number and mass of ^{4}He are given by $N_\mathrm{He} \approx N_\mathrm{n}/2$ and $M_\mathrm{He} \sim 4M_\mathrm{N}$, respectively. A combination of Eqs. (12.4.5) and (12.4.8) leads us to

$$Y_\mathrm{p} \approx \frac{2N_\mathrm{n}}{N_\mathrm{p} + N_\mathrm{n}} \approx 0.246 \tag{12.4.9}$$

It is clear that Y_p measures the abundance of primordial ^{4}He originating from the BBN process in the early Universe. To determine the value of Y_p as precisely as possible, one needs to measure the ^{4}He abundance in regions of hot ionized gas from "metal-poor" galaxies (e.g. the HII region in dwarf galaxies) where the chance of producing heavier elements than ^{4}He via the stellar burning of hydrogen is much smaller.

The observed values of Y_p are $Y_\mathrm{p} = 0.249 \pm 0.009$ [19]; $Y_\mathrm{p} = 0.248 \pm 0.003$ [20]; $Y_\mathrm{p} = 0.254 \pm 0.003$ or 0.252 ± 0.001 [21]; $Y_\mathrm{p} = 0.2534 \pm 0.0083$ [22]; and $Y_\mathrm{p} = 0.2465 \pm 0.0097$ [23]. These measurements are consistent with the prediction of the BBN theory in Eq. (12.4.9).

A more careful analysis yields [24]

$$Y_\mathrm{p} \sim 0.248 + 0.0096 \times \ln\left(\frac{\eta}{6.15 \times 10^{-10}}\right) + 0.013 \times (N_\nu - 3) \tag{12.4.10}$$

where η is the cosmic baryon-to-photon ratio n_B/n_γ, and N_ν denotes the number of light neutrino species. N_ν is apparently related to g_* in Eq. (12.2.23), and thus it may affect the BBN process. So the measured value of Y_p may reversely allow us to determine the value of N_ν. Given $N_\nu = 3$ extracted from the precision measurements at the LEP $\mathrm{e}^+\mathrm{e}^-$

collider, it is possible to determine the value of $Y_{\rm p}$ to a higher degree of accuracy by means of Eq. (12.4.10).

After the synthesis of ^{4}He, it is difficult for much heavier nuclei to form in any significant quantity because of both the absence of stable nuclei with the mass number $A = 5, 6, 8$ and the large Coulomb barriers which make the synthesis of ^{12}C from three ^{4}He strongly suppressed. But ^{7}Li and ^{7}Be can be produced via the following reactions:

$$\begin{aligned} {}^3{\rm H} + {}^4{\rm He} \to {}^7{\rm Li} + \gamma, \quad {}^3{\rm He} + {}^4{\rm He} \to {}^7{\rm Be} + \gamma \\ {}^7{\rm Be} + {\rm e}^- \to {}^7{\rm Li} + \nu_{\rm e}, \quad {}^7{\rm Be} + {\rm n} \to {}^7{\rm Li} + {\rm p} \end{aligned} \tag{12.4.11}$$

Here we do not elaborate on these processes, but refer the reader to Ref. [16] and references therein. Figure 12.2 shows the BBN predictions for the

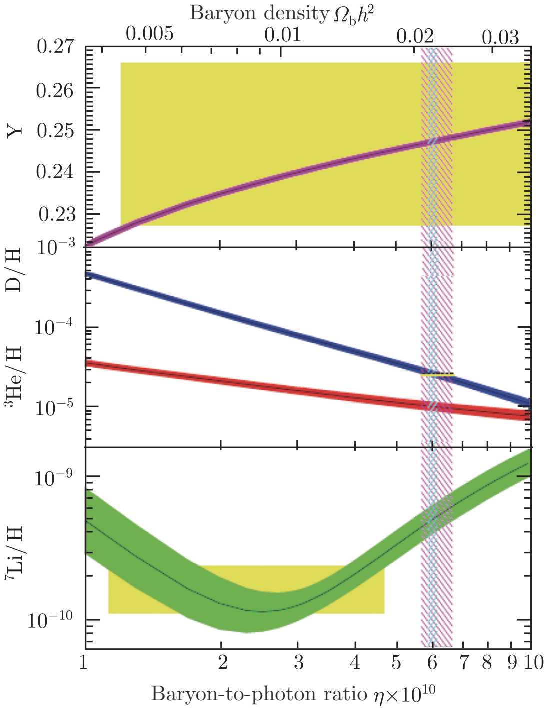


Fig. 12.2. The abundances of ^{4}He, D, ^{3}He and ^{7}Li predicted by the standard ΛCDM model. Here the bands show the allowed ranges at the 95% CL, and the upper and lower boxes indicate the observational data. The narrow vertical band stands for the CMB measurement of the cosmic baryon-to-photon ratio η, and the lower wider curve represents the BBN concordance range (both at the 95% CL).

abundances of light elements as compared with the corresponding experimental data.

12.5. Cosmological Baryon–Antibaryon Asymmetry

The observed baryon–antibaryon asymmetry of the Universe is also referred to as the matter–antimatter asymmetry. It is a long-standing puzzle in cosmology. In the past decades, a lot of efforts have been made to directly search for primordial antimatter on the Earth and in cosmic rays, the Milky Way Galaxy and outer spaces, but nothing has been found. The indirect search is also very useful in this connection. If there exists antimatter in space, matter–antimatter annihilations will happen and produce a huge amount of π^0 and $\pi^\pm$ mesons in their contact areas. A great amount of diffuse γ rays can be produced from the $\pi^0 \to 2\gamma$ decay, and their number density is peaked around 100–200 MeV. But a measurement of the diffuse gamma rays can only set an upper bound on antimatter, simply because the Compton scattering of the starlight and cosmic ray electrons may also contribute. In addition, the BBN and CMB measurements have given the values of $\eta = n_{\rm B}/n_\gamma$. One may obtain $\eta \sim \times 10^{-10}$ from Eq. (12.4.10) by inputting the observed value of $Y_{\rm p}$ for the primordial abundance of ^{4}He. The five-year WMAP result for the acoustic peaks in the angular power spectrum of CMB radiation leads us to [25]

$$\eta \equiv \frac{n_{\rm B} - n_{\bar{\rm B}}}{n_\gamma} = (6.21 \pm 0.16) \times 10^{-10} \tag{12.5.1}$$

Note that $n_{\bar{\rm B}} \approx 0$ holds today, so the value of η in Eq. (12.5.1) is consistent with that extracted from the BBN data. Although the Big Bang cosmology is successful in explaining the abundance of ^{4}He and the existence of CMB radiation, it is unable to explain why $\eta = n_{\rm B}/n_\gamma \sim 6 \times 10^{-10}$ and $n_{\bar{\rm B}} \approx 0$ hold. Even though $n_{\rm B} \neq n_{\bar{\rm B}}$ were assumed in the beginning of Big Bang, this asymmetry would not survive and would soon be washed out by cosmic inflation. So it is necessary to find out a dynamic mechanism to produce the cosmic baryon–antibaryon asymmetry. Such a baryogenesis mechanism should satisfy three prerequisites as outlined by Sakharov [26]:

(1) The existence of baryon-number-violating interactions;
(2) C and CP violation;
(3) Departure from thermal equilibrium.

The first requirement is satisfied by the standard model of particle physics, in which both baryon number B and lepton number L are accidently conserved at the classical level. The non-perturbative instanton effects violate B and L, but B–L is exactly conserved. At the electroweak scale, this kind of B-violating interaction is unable to produce a sufficiently large baryon–antibaryon asymmetry. In some GUTs, the quark decays mediated by superheavy particles X and Y are possible to give rise to sufficiently strong B violation (see the discussions on proton decays in Sec. 11.1).

The second condition requires simultaneous violation of C and CP symmetries. If C were conserved, the probabilities of producing baryons and antibaryons would be exactly equal, and thus the net baryon number would vanish due to an exact cancellation between an amount of B and the same amount of $-$B. That is why $B \neq 0$ definitely requires C violation. Now let us assume a superheavy particle X to decay through two channels with branching ratios r and $1 - r$, corresponding to baryon numbers B_1 and B_2. Then the CP-conjugate state $\bar{\mathrm{X}}$ decays into the CP-conjugate final states with branching ratios $\bar{r}$ and $1 - \bar{r}$, corresponding to baryon numbers $-B_1$ and $-B_2$. The CPT symmetry requires the numbers of X and $\bar{\mathrm{X}}$ to be equal, and hence decays of the $\mathrm{X}\bar{\mathrm{X}}$ pairs may result in a B-violating asymmetry

$$A_{\mathrm{B}\bar{\mathrm{B}}} = rB_1 - \bar{r}B_1 + (1 - r)B_2 - (1 - \bar{r})B_2 = (r - \bar{r})(B_1 - B_2) \tag{12.5.2}$$

B violation allows us to choose the decay modes which assure $B_1 \neq B_2$ to hold, and CP violation requires $r \neq \bar{r}$ to hold. That is why $A_{\mathrm{B}\bar{\mathrm{B}}} \neq 0$ requires not only C violation but also CP violation.

The third Sakharov condition can be understood as follows. If the probabilities of a physical reaction and its CP-conjugate process are smaller than the Hubble expansion rate of the Universe, they must depart from thermal equilibrium. The Universe staying in equilibrium can be regarded as an isolated system. In this case the ensemble average of baryon number is given by

$$\langle B \rangle = Z^{-1} \mathrm{Tr} \left[B \mathrm{e}^{-\beta H} \right] \tag{12.5.3}$$

where $Z \equiv \mathrm{Tr} \left[\mathrm{e}^{-\beta H} \right]$ is the partition function, $\beta = kT$, and H denotes the Hamiltonian. Assuming the system to be invariant under the CPT transformation, one has $(CPT)H(CPT)^{-1} = H$. Taking account of

$(CPT)B(CPT)^{-1} = -B$, one may rewrite Eq. (12.5.3) as

$$\begin{aligned}\langle B\rangle &= Z^{-1}\mathrm{Tr}\left[B\mathrm{e}^{-\beta H}\right] = Z^{-1}\mathrm{Tr}\left[(CPT)B(CPT)^{-1}(CPT)\mathrm{e}^{-\beta H}(CPT)^{-1}\right]\\ &= Z^{-1}\mathrm{Tr}\left[-B\mathrm{e}^{-\beta H}\right] = -\langle B\rangle \qquad (12.5.4)\end{aligned}$$

Then one is left with $\langle B\rangle = \langle -B\rangle = 0$. That is, the net baryon number must be vanishing if the Universe stays in thermal equilibrium. In short, the three Sakharov conditions have to be simultaneously satisfied so as to generate a net baryon–antibaryon asymmetry.

There are a number of viable models on the market which can explicitly realize dynamical generation of the cosmic baryon–antibaryon asymmetry. Here we simply outline a few typical ideas in this respect, and refer the reader to Chapter 11 of Ref. [16] and references therein.

The following baryogenesis mechanisms are currently popular:

(1) *Electroweak baryogenesis.* The standard electroweak theory satisfies all the three Sakharov conditions, but a successful realization of baryogenesis at the electroweak scale also requires the non-perturbative sphaleron effect not to be too strong and the electroweak phase transition to be of the first order. The latter is unfortunately unsatisfied within the standard model because the mass of the Higgs boson is too large.

(2) *GUT baryogenesis.* As shown in Sec. 11.1, the superheavy gauge bosons X and Y can both mediate the baryon-number-violating and lepton-number-violating interactions. In a GUT model *CP* violation is possible at the loop level, but it is also necessary to break (B–L) in order to make baryogenesis viable.

(3) *The Affleck-Dine mechanism.* In supersymmetric models, the superpartners of leptons and quarks are the scalar fields with both baryon and lepton numbers. It is therefore possible to introduce the B-violating and CP-violating terms in the corresponding scalar potential. The decays of scalar fields can convert baryon number to the standard model particles, so as to generate the cosmic baryon-antibaryon asymmetry in a dynamical way.

(4) *Leptogenesis.* The mechanism of baryogenesis via leptogenesis is currently the most popular way to account for the observed baryon–antibaryon asymmetry of the Universe, because it is closely associated with the seesaw mechanism for the origin of tiny neutrino masses. The key idea of this mechanism is to convert the lepton–antilepton asymmetry produced from the CP-violating and out-of-equilibrium decays

of heavy Majorana neutrinos in the early Universe to the baryon–antibaryon asymmetry via the non-perturbative sphaleron interactions in thermal equilibrium.

12.6. Dark Matter

The most robust evidence for the existence of dark matter in the Universe is from observations of the rotation curves of luminous galaxies. For the motion of a planet in the solar system, its tangential velocity is related to the Sun's gravitation as follows:

$$\frac{mv^2}{r} = G_{\rm N}\frac{mM}{r^2} \tag{12.6.1}$$

where m is the mass of this planet, M is the mass of the Sun, r is the distance between the planet and the Sun, and the planetary orbit has been taken to be circular as a good approximation. Then we obtain

$$v = \sqrt{G_{\rm N}M/r} \propto \frac{1}{\sqrt{r}} \tag{12.6.2}$$

The observational data on the motion of nine planets in the solar system fully support the relationship given in Eq. (12.6.2).

But the motion of stars rotating in a galaxy does not satisfy the above relationship at all. Figure 12.3 shows the rotation curves of the spiral galaxy NGC6503 (i.e. the $v - r$ curves) [18]. One can see that the tangential velocities of stars in this galaxy are almost flat against the radial distance r; namely, $v \sim$ is roughly a constant for $r > 4\,$kpc. This observation tells us that the galactic mass is not only from the luminous material at the galactic center but also from the non-luminous halo, and the latter constitutes the bulk of the galactic mass.

To obtain $v \sim$ constant, the galactic mass density ρ must satisfy

$$\rho \propto r^{-2} \tag{12.6.3}$$

and the non-luminous halo dominates the total galactic mass

$$M \approx \int_0^r \rho(r){\rm d}^3r \propto r \tag{12.6.4}$$

Substituting Eq. (12.6.4) into Eq. (12.6.1), one gets

$$\frac{mv^2}{r} = \frac{G_{\rm N}m\int_0^r \rho(r){\rm d}^3r}{r^2} \propto \frac{1}{r} \tag{12.6.5}$$

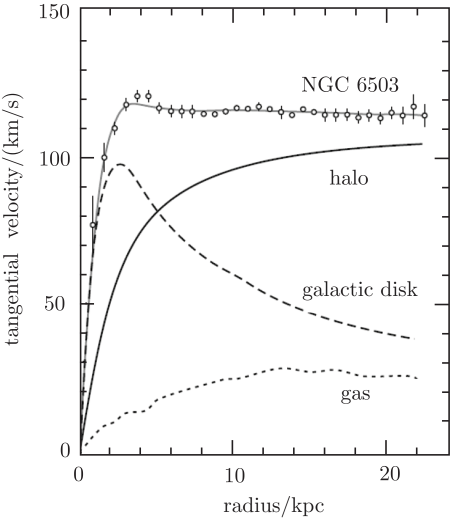


Fig. 12.3. The tangential velocities of stars in the spiral galaxy NGC6503 changing with the radial distance r, where the contributions from the disc, gas and halo are separately shown.

where $M(r) \approx \int_0^r \rho \mathrm{d}^3x \propto r$, proportional to the radial distance r. As a result,

$$v = \text{constant} \tag{12.6.6}$$

The phenomena like that shown in Fig. 12.3 exist in those much larger clusters and superclusters of galaxies, indicating the existence of non-luminous matter — dark matter in the Universe.

What is dark matter made of? Is it baryonic matter? The answers to these two questions remain unclear to us. But the observational data listed in Eq. (12.1.26) tell us that cold dark matter constitutes about 27% of the energy density of today's Universe, and the contribution of baryonic matter is only about 5%. Note that only about $\leq 1\%$ of baryonic matter is luminous (visible), and the other 4% is non-luminous (invisible), much less than the total amount of dark matter. A few possible candidates for dark matter are listed below.

(1) **MACHOs**: In the Milky Way galaxy a massive compact halo object (MACHO) has been observed from gravitational lensing. Suppose that there is a MACHO between a luminous star and the observer. When

one observes the star, the gravitational deflection of light caused by the MACHO may form a ring of light around the star — the so-called Einstein ring. Namely, the observer will see two images of the source: a luminous star and a ring of light. But if the MACHO's mass $M_{\rm d}$ is not big enough, the Einstein ring will not be big enough to make the two overlapping images distinguishable. In this case, the gravitational lensing effect allows light from the star to be focused, such that the luminosity of the star increases. This kind of luminosity enhancement is limited to a given period, and it will disappear when the MACHO is not collinear with the source and the observer. This period is given by $t \propto r_{\rm E}/v$, where $r_{\rm E} \sim \sqrt{M_{\rm d}}$ ($M_{\rm d}$ is the dark halo mass) is the radius of the Einstein ring and v is the velocity of the MACHO. Figure 12.4 illustrates the luminosity enhancement of light from a star in the Large Magellanic Cloud [27]. Because of $\Delta p \propto p$ for the change of the photon momentum caused by a gravitational field, the deflection of light $\Delta p/p$ is independent of its wavelength (or frequency). That is why, the peaks of blue and red light signals are identical in Fig. 12.4.

The observational results shown in Fig. 12.4 indicate that there are really MACHOs in the Large Magellanic Cloud, and they also exist in other clusters of galaxies. A careful analysis tells us that they belong

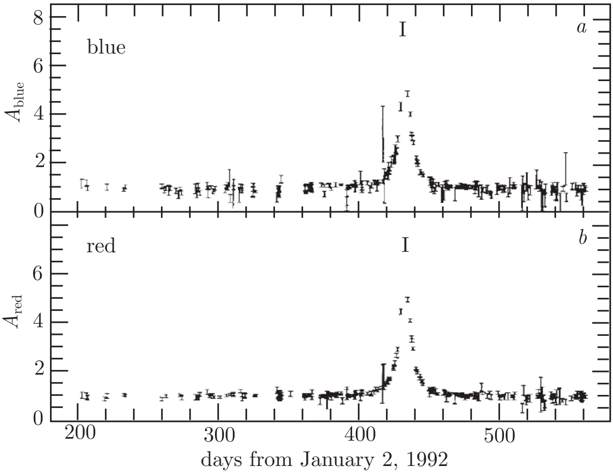


Fig. 12.4. The luminosity enhancement of light from a star in the Large Magellanic Cloud caused by the gravitational lensing effect. The horizontal axis denotes time in days, and the vertical axis stands for an amplification of the light signal, where $A_{\rm red}$ and $A_{\rm blue}$ correspond to red and blue light signals, respectively.

to baryonic dark matter. A few possible candidates for non-baryonic dark matter will be introduced below.

(2) **Neutrinos**: The weakly interacting neutrinos ($\nu_{\rm e}$, ν_μ and ν_τ) and their antiparticles can only serve as "hot dark matter", because they are relativistic particles at the onset of galaxy formation when the temperature of the Universe is about 1 keV. Today, their energy density in the Universe is very small, $\Omega_\nu = \rho_\nu/\rho_c < 0.0055$, as one can see from Eq. (12.1.26). That is why, the active neutrinos do not constitute the major part of dark matter.

(3) **Axions**: The axion is a hypothetical pseudoscalar Goldstone particle which can help to solve the strong CP problem, and its mass is expected to be around $10^{-6} \sim 1$ eV. If axions are non-thermally produced in the early Universe, they may serve as "cold dark matter".

(4) **WIMPs:** The weakly interacting massive particles (WIMPs) are currently the most popular candidate for cold dark matter in the direct-search experiments. A lot of attention has been paid to the lightest stable particles (LSPs), such as neutrolinos predicted by many supersymmetric models. So far a lot of international efforts have been made in searching for this kind of hypothetical particles both directly (in underground experiments) and indirectly (in astronomical observations), but no convincing signals have been found. (For a review with extensive references, we refer the reader to the PDG review of Ref. [1].)

12.7. Dark Energy

So far all the astronomical observations show that the standard ΛCDM model is correct. For the expansion of the Universe, one may define a deceleration parameter q_0 as follows:

$$q_0 = -\left.\frac{R\ddot{R}}{\dot{R}^2}\right|_0 = \frac{1}{2}\Omega_{\rm m} + \Omega_{\rm r} + \frac{(1+3w)}{2}\Omega_{\rm V} \qquad (12.7.1)$$

where the subscript "0" denotes the present time. Given Eq. (12.1.26) with $\Omega_{\rm r} \ll 1$, $\Omega_{\rm m} \sim 0.315$, $\Omega_{\rm V} = \Omega_\Lambda \sim 0.685$ and $w = -1$ in the ΛCDM model, we find

$$q_0 \approx \frac{1}{2}\Omega_{\rm m} - \Omega_\Lambda \approx -0.53 < 0 \qquad (12.7.2)$$

implying that the expansion of the Universe is accelerating; namely,

$$\ddot{R}(t_0) > 0 \qquad (12.7.3)$$

In fact, Eq. (12.1.26) tells us that today's Universe is dominated by vacuum energy. Given $R(t) = \mathrm{e}^{\sqrt{\frac{\Lambda}{3}}t}$ in Eq. (12.1.50), it is straightforward to obtain

$$\ddot{R}(t) = \frac{\Lambda}{3}\mathrm{e}^{\sqrt{\frac{\Lambda}{3}}t} > 0 \tag{12.7.4}$$

indicating that the Hubble expansion is accelerating. This observation has been verified by the experimental data on Type Ia supernova luminosities [28]. Figure 12.5 shows the observational results, from which one can see the redshift of light from Type Ia supernova explosion is larger than the prediction based on a linear Hubble's law. The best fit to all available data today yields $\Omega_{\mathrm{m}} \sim 0.3$ and $\Omega_{\mathrm{V}} \sim 0.7$ in the neglect of Ω_{r} and in the assumption of $\kappa = 0$ within the ΛCDM model. Therefore, we conclude that the expansion of our Universe is accelerating.

Obviously, the cosmological constant Λ and the ΛCDM model can provide the simplest theoretical explanation of the Universe's accelerating expansion. But Λ leads us to the vacuum energy density $\rho_\Lambda = \frac{\Lambda}{8\pi G_{\mathrm{N}}} \sim (2.2\,\mathrm{meV})^4 \sim 10^{-123} M_{\mathrm{P}}^4$, where meV$= 10^{-3}$ eV. Such a tiny vacuum energy density looks quite unnatural from a theoretical point of view. It is very hard to understand this result, because the vacuum energy in any quantum field theory is expected to be ≥ 1 TeV [29] instead of meV, and these two values span 15 orders of magnitude. A common conjecture is therefore that the vacuum energy might not be a constant; instead, it might change with time. The reason for the smallness of vacuum energy today might be that our Universe is so old that the vacuum energy becomes smaller and smaller for such a long time after the Big Bang. The time-varying vacuum energy is also called *quintessence* [30]. The simplest way to achieve time-varying vacuum energy is to introduce a new type of scalar field. For example, assuming the energy–momentum tensor of a scalar field ϕ to be

$$T_{\mu\nu} = \partial_\mu\phi\partial_\nu\phi - g_{\mu\nu}\partial_\rho\phi\partial^\rho\phi - g_{\mu\nu}V(\phi) \tag{12.7.5}$$

and using $\rho_\phi = T_{00}$ and $p_\phi = R^{-2}(t)T_{ii}$, we find that its energy density ρ_ϕ and pressure p_ϕ are given by

$$\begin{cases} \rho_\phi = \dfrac{1}{2}\dot{\phi}^2 + \dfrac{1}{2}R^{-2}(t)(\nabla\phi)^2 + V(\phi) \\[2ex] p_\phi = \dfrac{1}{2}\dot{\phi}^2 - \dfrac{1}{6}R^{-2}(t)(\nabla\phi)^2 - V(\phi) \end{cases} \tag{12.7.6}$$

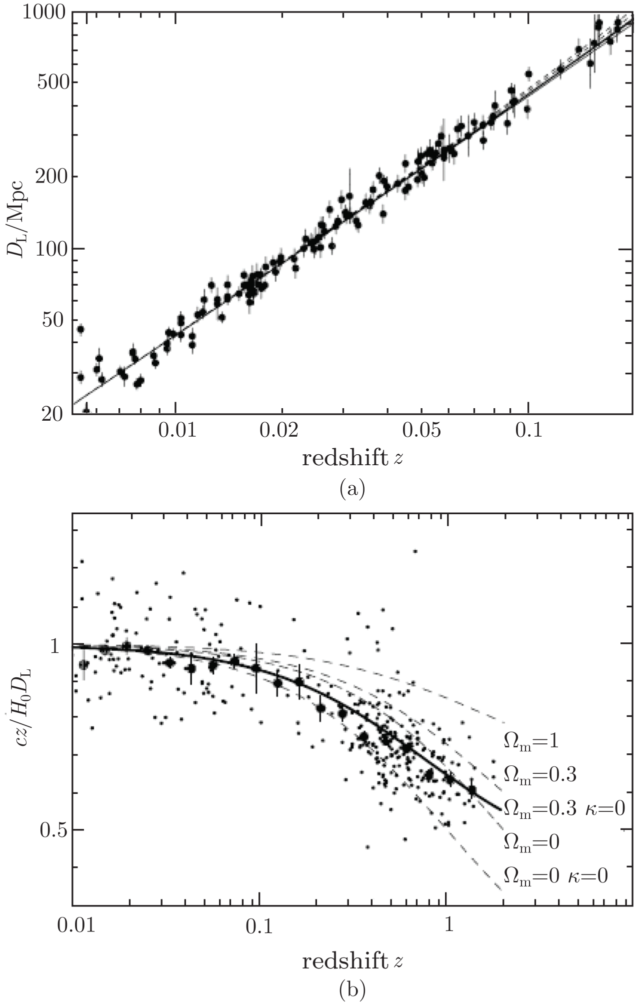


Fig. 12.5. The Hubble plot for Type Ia supernova explosion. (a) The linearity of Hubble's law at $z \ll 1$; (b) The result at $z > 0.1$, where the linearity of Hubble's law has been broken. The behavior of the curves indicates that the redshift of light from Type Ia supernova explosion is larger than the prediction based on a linear Hubble's law, where D_L is the luminosity distance between the supernova and the observer on the Earth.

In a homogeneous space, the $\nabla\phi$ term is negligible, and thus the equation of motion of ϕ reads

$$\ddot{\phi} + 3H\dot{\phi} = -\partial V/\partial\phi \tag{12.7.7}$$

In this case, Eq. (12.7.6) can be simplified to

$$\begin{cases} \rho_\phi = \dfrac{1}{2}\dot{\phi}^2 + V(\phi) \\[2ex] p_\phi = \dfrac{1}{2}\dot{\phi}^2 - V(\phi) \end{cases} \tag{12.7.8}$$

If $\frac{1}{2}\dot{\phi}^2 \ll |V(\phi)|$ holds, we will be left with $w_\phi = p_\phi/\rho_\phi = -1$. In general, $\frac{1}{2}\dot{\phi}^2$ cannot be neglected, so Eq. (12.7.8) leads us to $(1 + w_\phi)\rho_\phi = \dot{\phi}^2 > 0$. If and only if $\rho_\phi > 0$, we must arrive at

$$w_\phi = p/\rho > -1 \tag{12.7.9}$$

implying that the energy density of ϕ may act as the cosmological constant and thus make the Hubble expansion accelerate in the later period of the Universe. For instance, consider

$$V(\phi) = M^{4+\alpha}\phi^{-\alpha} \tag{12.7.10}$$

where $\alpha > 0$ is an arbitrary real number and M can be regarded as an effective mass. In the early Universe dominated by radiation, $R(t) \propto \sqrt{t}$ and $H = \frac{1}{2t}$ hold, as shown in Eq. (12.1.45). Substituting Eq. (12.7.10) into Eq. (12.7.7), we obtain

$$\ddot{\phi} + \frac{3}{2t}\dot{\phi} - \alpha M^{4+\alpha}\phi^{-\alpha-1} = 0 \tag{12.7.11}$$

Solving this equation, we find

$$\phi = \left[\frac{\alpha(2+\alpha)^2 M^{4+\alpha}}{6+\alpha} t^2\right]^{\frac{1}{2+\alpha}} \tag{12.7.12}$$

It is clear that both $\dot{\phi}^2$ and $V(\phi)$ are proportional to $t^{-\frac{2\alpha}{2+\alpha}}$, and hence $\rho_\phi \propto t^{-\frac{2\alpha}{2+\alpha}}$. When t is very small in the early Universe, $\rho_{\rm r} \propto R^{-4} \sim t^{-2}$ holds, and therefore $\rho_\phi < \rho_{\rm r}$ can be obtained. This result is consistent with the fact that the early Universe is dominated by radiation.

When the energy density of radiation becomes lower than the energy density of non-relativistic matter in the Universe, Eq. (12.1.47) tells us that $R \propto t^{2/3}$ and $H = \frac{2}{3t}$ hold. In this case, the coefficient of the second term on the left-hand side of Eq. (12.7.11) is replaced as $\frac{3}{2t} \to \frac{2}{t}$, so the solution of ϕ in Eq. (12.7.12) will have a different coefficient but its relation with t keeps unchanged. In other words, ϕ remains to increase in proportion to $t^{\frac{2}{2+\alpha}}$, while $\dot{\phi}^2$, $V(\phi)$ and ρ_ϕ remain to decrease in proportion to $t^{-\frac{2\alpha}{2+\alpha}}$. But now $\rho_{\rm r}$ and $\rho_{\rm m}$ decrease faster in proportion to $t^{-8/3}$ (i.e. R^{-4}) and t^{-2} (i.e.

R^{-3}), respectively. Therefore, it is always possible that $\rho_{\rm m}$ and $\rho_{\rm r}$ become smaller than ρ_ϕ at some moment. In short, we have found a solution to the scalar potential, whose energy density decreases with time. So it may serve as the cosmological constant in the later period of the Universe. For some further discussion, we refer the reader to Sec. 1.12 of Ref. [8].

One may also modify Newton's theory of gravitation to interpret the accelerating expansion of the Universe, but there are many obstacles in doing so. For more discussions in this connection, we refer the reader to the relevant reviews included in Ref. [1].

12.8. Cosmic Microwave Background

The cosmic microwave background (CMB) radiation was first observed in 1965 [14], and its black-body radiation spectrum and tiny anisotropy were precisely measured by the COBE satellite [31]. The CMB anisotropy can be used to strictly constrain many of the cosmological parameters. In particular, the Wilkinson Microwave Anisotropy Probe (WMAP) satellite has reported more accurate data on the CMB anisotropy, pushing cosmology into the era of precision cosmology.

12.8.1. *Transition of the Universe to its matter-dominated epoch*

As the Universe expands, its energy density will gradually transform from the radiation-dominated epoch (with photons and relativistic particles) to the matter-dominated epoch (with non-relativistic particles). Afterwards, the photons become decoupled from matter and form the CMB radiation as observed today.

First of all, let us take a look at when the energy density of radiation (photons and neutrinos) becomes equal to that of matter.

For a flat Universe with $\kappa = 0$, Eq. (12.1.26) provides us with today's observational data $\Omega_{\rm m}(t_0) \approx 0.315$, $\Omega_\Lambda = \Omega_{\rm V} \approx 0.685$, $\Omega_\gamma(t_0) \approx 5.46 \times 10^{-5}$ and $\Omega_{\rm b}(t_0) \approx 0.0499$. With the help of Eq. (12.3.9), we obtain

$$\frac{\rho_\nu}{\rho_\gamma} = \frac{7}{8}\frac{g_\nu}{g_\gamma}\left(\frac{T_\nu}{T_\gamma}\right)^4 = \frac{21}{8}\left(\frac{4}{11}\right)^{4/3} \approx 0.681 \qquad (12.8.1)$$

As a result,

$$\Omega_\nu(t_0) = \frac{\rho_\nu}{\rho_\gamma}\Omega_\gamma(t_0) \approx 3.72 \times 10^{-5} \qquad (12.8.2)$$

where we have assumed all the neutrinos to be massless. Then we have

$$\Omega_{\mathrm{r}}(t_0) = \Omega_\gamma(t_0) + \Omega_\nu(t_0) = 9.18 \times 10^{-5} \tag{12.8.3}$$

Note that $\Omega_\gamma \propto \rho_\gamma \propto R^{-4}$, $\Omega_{\mathrm{m}} \propto \rho_{\mathrm{m}} \propto R^{-3}$, $\Omega_{\mathrm{r}} \propto \rho_{\mathrm{r}} \propto R^{-4}$ and $\Omega_{\mathrm{b}} \propto \rho_{\mathrm{b}} \propto R^{-3}$ hold, and thus from Eq.(12.1.26)

$$\frac{\Omega_\gamma(t)}{\Omega_{\mathrm{m}}(t)} = \frac{\Omega_\gamma(t_0)R^4(t_0)}{R^4(t)} \cdot \frac{R^3(t)}{R^3(t_0)\Omega_{\mathrm{m}}(t_0)} = \frac{\Omega_\gamma(t_0)}{\Omega_{\mathrm{m}}(t_0)}(1+z) \tag{12.8.4}$$

where $\frac{R(t_0)}{R(t)} = 1 + z$ in Eq. (12.1.37) has been used. So the time for the matter-radiation equality satisfies

$$\frac{\Omega_\gamma(t)}{\Omega_{\mathrm{m}}(t)} = (1+z)\frac{\Omega_\gamma(t_0)}{\Omega_{\mathrm{m}}(t_0)} = 1 \tag{12.8.5}$$

from which the corresponding redshift can be derived:

$$z_{\mathrm{m}\gamma} = \frac{\Omega_{\mathrm{m}}(t_0)}{\Omega_\gamma(t_0)} - 1 \approx 5769 \tag{12.8.6}$$

In a similar way, one may derive the redshift corresponding to the photon–baryon equality:

$$z_{\mathrm{b}\gamma} = \frac{\Omega_{\mathrm{b}}(t_0)}{\Omega_\gamma(t_0)} - 1 \approx 914 \tag{12.8.7}$$

The redshift corresponding to the equality between the energy density of the radiation (photons and neutrinos) and that of matter is given by

$$z_{\gamma\mathrm{m}} = \frac{\Omega_{\mathrm{m}}(t_0)}{\Omega_\gamma(t_0) + \Omega_\nu(t_0)} - 1 \approx 3431 \tag{12.8.8}$$

To figure out the time corresponding to the matter-radiation equality in Eq. (12.8.6) and that corresponding to the matter-radiation (photon+neutrino) equality in Eq. (12.8.8), we assume that the Universe has been matter-dominated since the matter-radiation equality. In this case $t \propto R(t)^{3/2}$ holds as shown in Eq. (12.1.47), and thus we arrive at

$$t = t_0 \left[\frac{R(t)}{R(t_0)}\right]^{3/2} = \frac{t_0}{(1+z)^{3/2}} \tag{12.8.9}$$

Substituting $t_0 = 13.81\,\text{Gyr}$ into the above equation, one may obtain the time for the matter-radiation (photon+neutrino) equality ($z_{\gamma\text{m}} = 3431$) as

$$t_{\gamma\text{m}} \approx 6.9 \times 10^4\text{yr} \tag{12.8.10}$$

and the time for the matter-photon equality ($z_{\text{m}\gamma} = 5769$) as

$$t_{\text{m}\gamma} \approx 3.2 \times 10^4\text{yr} \tag{12.8.11}$$

It should be noted that these values are far from accurate. Moreover, we have not considered the fact that vacuum energy actually dominates the energy density of the Universe in its later period with $t \propto \ln R(t)$.

12.8.2. *Formation of the CMB*

In the early Universe photons are so energetic that any electrons and protons managing to bind into neutral hydrogen atoms via the reaction $\text{e} + \text{p} \to \text{H} + \gamma$ may easily be photodissociated by the inverse process $\gamma + \text{H} \to \text{p} + \text{e}$. Namely, the Universe exhibits an ionized plasma of electrons and protons. As the temperature drops continuously, the photon energy becomes smaller and smaller, making the probability of $\gamma + \text{H} \to \text{p} + \text{e}$ strongly suppressed. In this case, the Universe transforms from the e–p plasma into a gas of neutral hydrogen atoms. This process is called "recombination". After recombination, the number density of free electrons in the Universe is so low that the mean free path of a photon becomes too long to be scattered with matter anymore. In other words, photons become decoupled from matter.

What is the time of recombination? A naive expectation is that the inverse process of the reaction $\text{e} + \text{p} \Leftrightarrow \text{H} + \gamma$ should not take place if the temperature of the Universe is below the binding energy of hydrogen 13.6 eV. But because of $\eta = \frac{n_\text{B}}{n_\gamma} \sim 6 \times 10^{-10}$, it is very easy for a small fraction of photons to have energies above 13.6 eV. So the temperature T is required to far below 13.6 eV in order to assure that the number of photons with energies above 13.6 eV is smaller than the number of baryons (hydrogen atoms). A careful analysis [32] yields the temperature of recombination $T_{\text{rec.}} \sim 0.3\,\text{eV}(\sim 3500\,\text{K})$. Now let us calculate the time of recombination $t_{\text{rec.}}$ corresponding to $T_{\text{rec.}}$. Given $T \propto R^{-1}$, Eq. (12.1.37) leads us to

$$\frac{T_{\text{rec.}}}{T_0} = \frac{R_0}{R_{\text{rec.}}} = 1 + z \tag{12.8.12}$$

Inputting $T_0 \approx$2.73K, we obtain $T_{\text{rec.}} = \frac{0.3\,\text{eV}}{k} = \frac{0.3\,\text{eV}}{8.617\times 10^{-11}\,\text{MeV K}^{-1}} =$ 3481 K and

$$1 + z_{\text{rec.}} = \frac{T_{\text{rec.}}}{T_0} = \frac{3481}{2.73} \approx 1275 \tag{12.8.13}$$

Assume the evolution of the Universe to be dominated by matter, $t \sim R(t)^{3/2}$. Then we make use of Eq. (12.8.9) and arrive at

$$\begin{aligned} t_{\text{rec.}} &= t_0 \left[\frac{R(t)}{R(t_0)}\right]^{3/2} = t_0 \left[\frac{T_0}{T_{\text{rec.}}}\right]^{3/2} \\ &= \frac{t_0}{(1+z_{\text{rec.}})^{3/2}} \\ &\approx 303561\,\text{yr} \\ &\approx 3.0 \times 10^5\,\text{yr} \end{aligned} \tag{12.8.14}$$

Note that photons are not immediately decoupled from matter after recombination. In fact, the decoupling of photons from matter need to take another little time. The mean free path of photons can be calculated from the cross section of the Thomson scattering process and the number density of free electrons. When such a path length is larger than the horizon distance (i.e. the radius of the Universe observed at a given time), photons will be decoupled from matter. This decoupling temperature $T_{\text{dec.}} \sim 0.26$ eV is equivalent to 3017 K, corresponding to $1 + z_{\text{dec.}} \approx 1104$. The decoupling time turns out to be

$$t_{\text{rec.}} = t_0 \left(\frac{T_0}{T_{\text{dec.}}}\right)^{3/2} = \frac{t_0}{(1+z_{\text{dec.}})^{3/2}} \approx 376754\text{yr} \approx 3.8 \times 10^5\text{yr} \tag{12.8.15}$$

After the decoupling of photons from matter, they can freely travel in the Universe without any obstacles until today. One may define a "surface of last scattering" as the sphere centered about us with a radius equal to the mean distance to the last place where the CMB photons have scattered with matter. The CMB radiation observed today actually reflects the situation of the Universe at a time of about 3.8×10^5 years after the Big Bang.

12.8.3. *Anisotropies of the CMB*

As early as in 1965, Penzias and Wilson observed the homogeneous and isotropic CMB radiation of about 3K in the Universe [14]. At that time, the error bar was quite large, around 10%. An anisotropy of the CMB radiation at the 10^{-3} level was first measured in the 1970s. It is a reflection of the fact that the CMB temperature is a little higher in one particular direction of the microwave sky than in the opposite direction. This anisotropy is caused by the Doppler effect coming from the Earth's motion relative to the CMB (equivalent to the velocity of the Milky Way galaxy relative to the CMB, $v_{\rm MW} \sim 630\,{\rm km\ s}^{-1}$).

In 1992, the COBE satellite observed the CMB temperature variations at smaller angular separations [31]. Such tiny temperature variations, which are only at the 10^{-5} level, have been verified by many experimental collaborations. In particular, the WMAP satellite made a series of measurements of the CMB anisotropies during 2003 to 2009, and its sensitivity to angular separations was as precise as 0.1° [33].

To describe the CMB anisotropies, one may consider the CMB temperature as a direction-dependent function $T(\theta, \varphi)$, where θ and φ are the polar and azimuthal angles in a spherical coordinate system. This function can be expanded in terms of the spherical harmonic functions $Y_{lm}(\theta, \varphi)$ as

$$T(\theta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} a_{lm} Y_{lm}(\theta, \varphi) \qquad (12.8.16)$$

where the larger values of l correspond to the smaller angular scales, $l = 0$ corresponds to the monopole term, $l = 1$ corresponds to the dipole term, and so on.

Define

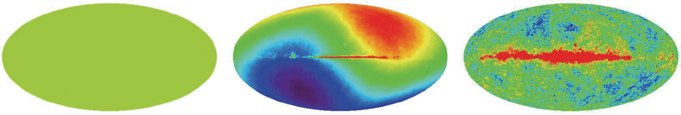
$$c_l = \frac{1}{2l+1} \sum_{m=-l}^{l} |a_{lm}|^2 \qquad (12.8.17)$$

where c_l is called the angular power spectrum, and it describes the structure of the CMB with the angular separation

$$\Delta\theta = \frac{180^\circ}{l} \qquad (12.8.18)$$

In a realistic experiment only some of the values of l can be measured because the experimental sensitivity to $\Delta\theta$ is limited when l is very large.

(1) *The monopole.* The $l = 0$ term in the Laplace expansion of $T(\theta, \varphi)$ given by Eq. (12.8.16) is referred to as the monopole component of the CMB spectrum. It represents the temperature of the CMB radiation



(a) (b) (c)

Fig. 12.6. The Patterns in the cosmographic map of the CMB measured by the WMAP satellite: (a) the monopole; (b) the dipole; (c) the multipoles ($l \geq 2$) (Courtesy of the WMAP Science Team, http://wmap.gsfc.nasa.gov/).

averaged over all directions of the sky, and its value is $T_\gamma = 2.725 \pm 0.001\,\mathrm{K}$ as measured by the COBE satellite [34]. The monopole distribution of the CMB radiation is in good agreement with a black-body function with $T_\gamma = 2.725\mathrm{K}$, as shown in Fig. 12.6(a).

(2) *The dipole.* The dipole ($l = 1$) term has an amplitude $\Delta T = 3.355 \pm 0.008\mathrm{mK}$ [35], leading us to $\frac{\Delta T}{T_\gamma} = 1.23 \times 10^{-3}$. The dipole distribution of the CMB radiation is shown in Fig. 12.6(b), which exhibits a TaiJi "yin-yang" pattern. It represents the largest anisotropy of the CMB, which is attributed to the Doppler shift caused by the Earth's motion relative to the isotropic black-body radiation field. The dipole anisotropy means that the CMB is blueshifted to a slightly higher temperature in the direction of the Earth's motion and redshifted in the opposite direction. It is obvious that the CMB temperature fluctuation at the 10^{-3} level belongs to the secondary anisotropy effect instead of the primordial fluctuation. One may certainly imagine to observe the Universe in a "local rest frame" fixed to the isotropic black-body radiation field. In this case, the CMB radiation would have no dipole anisotropy at all.

(3) *The multipoles.* ($l \geq 2$) In Eq. (12.8.16), the $l \geq 2$ terms in the Laplace expansion of $T(\theta, \varphi)$ represent the smaller-angle anisotropies of the CMB spectrum, characterized by the temperature fluctuations $\Delta T/T_\gamma \sim 10^{-5}$. In 1992, the COBE satellite first measured such tiny CMB anisotropies with an angular resolution $\Delta\theta \sim 7^\circ$ [31]. The latter is equivalent to $l \sim \pi/\Delta\theta \sim 20$. The multipole pattern in the cosmographic map of the CMB radiation is shown in Fig. 12.6(c), where the dipole component has been subtracted. This pattern comes from the measurements done by the WMAP collaboration, and it has an angular resolution $\Delta\theta \sim 0.2^\circ$ (corresponding to $l \sim 1000$ [33]). The 10^{-5}-level CMB temperature fluctuations can be attributed to small density fluctuations in the early Universe, manifesting themselves in the epoch

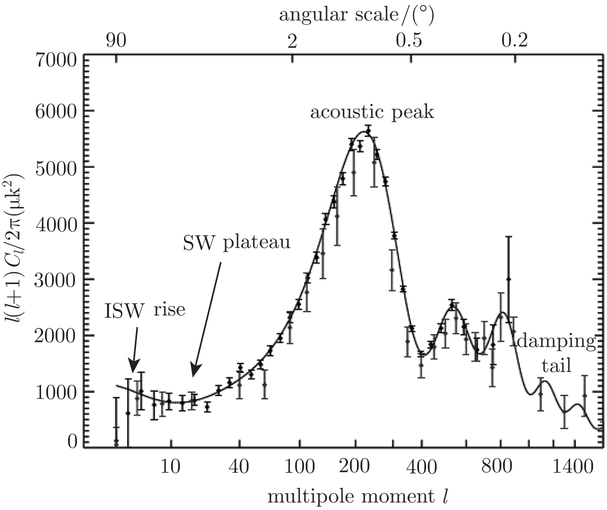


Fig. 12.7. The angular spectrum of the CMB. The points with smaller error bars are the values of the WMAP measurement, and those with larger error bars are from the previous measurements.

of last scattering of the CMB photons. Note that the density fluctuations in the early Universe can be amplified by gravity, with denser regions attracting more matter, until the non-relativistic matter of the Universe is separated into clumps (the formation of galaxies). The phenomena of cosmic density fluctuations should correspond to regions of different temperatures in the CMB radiation, which have been verified by the COBE and WMAP measurements [31, 33]. The angular power spectrum of the CMB is usually plotted as $l(l+1)c_l/(2\pi)$ against the multipole number l, as shown in Fig. 12.7 [36]. This spectrum can be separated into three main regions.

(1) The ISW rise ($l \leq 10$) and the SW plateau ($10 \leq l \leq 100$).

The integrated Sachs–Wolfe (ISW) rise and the Sachs–Wolfe (SW) plateau belong to the secondary anisotropy effects. When the CMB photons travel from last scattering to the observer on the Earth, they have to pass through galaxies and dark matter on their way and thus are affected by a time-varying gravitational potential. The fluctuations of this gravitational potential will result in the CMB anisotropies.

(2) The acoustic peaks ($100 \leq l \leq 1000$). The acoustic peaks belong to the primordial anisotropy, arising from the oscillations of the

photon–baryon fluid before the decoupling of photons from matter. When the temperature of the Universe is higher than the decoupling temperature of photons, the baryonic matter (composed of ions and electrons) and photons constitute a plasma. The gravitation is stronger in the denser region of this plasma, and thus more matter will be attracted to this region, making the local temperature increase. The hot matter radiates photons, making the pressure of photons increase. So the plasma of baryons will expand, and then it will cool down and hence produce less radiation. With decreasing radiation pressure the irregularities may reach a point at which gravity again takes over and initiates another compression phase. A competition between the gravitational accretion and radiation pressure gives rise to longitudinal acoustic oscillations in the photon–baryon fluid. After the decoupling of matter from radiation, the pattern of acoustic oscillations can be frozen into the CMB. As a consequence, the CMB anisotropies will be produced.

The first acoustic peak appears at $l \sim 200$, indicating that the total energy density of the Universe Ω is very close to unity. The latter implies that the Universe is essentially flat. The other peaks allow us to determine the baryon density $\Omega_{\rm b}$ and the cold dark matter density $\Omega_{\rm cd}$. For a detailed discussion in this connection, we refer the reader to Ref. [32].

(3) The damping tail ($l \geq 1000$). The damping tail is caused by the disappearance of the photon–baryon fluid as photons are decoupled and hydrogen atoms are formed. The mean free path of photons increases quickly, making the CMB anisotropies damp quickly.

12.9. Three Cosmic Problems and the Inflation Mechanism

12.9.1. *Three problems in cosmology*

The Big Bang ΛCDM model proves to be greatly successful in describing the evolution of the Universe, but it has three difficulties in how to understand the initial conditions of this evolution. Namely, how to solve the horizon problem, the flatness problem and the magnetic monopole problem in cosmology. In the following, we summarize these issues and explain why the inflation mechanism can help to solve them.

12.9.1.1. *The horizon problem*

If the tiny anisotropies of the CMB (caused by the motion of the Earth) are removed, it will be homogeneous and isotropic to a high degree of accuracy ($\sim 10^{-5}$). The time for the decoupling of radiation from matter is $t_{\text{dec.}} \approx 3.8 \times 10^5$yr, by which the maximal diameter of the causally connected (or interaction) region in the Universe is about $2ct_{\text{dec.}}$. According to the Big Bang theory, the expansion of such a region to date ($t_0 \approx 1.4 \times 10^{10}$yr) should have a diameter

$$2ct_{\text{dec.}}\frac{R(t_0)}{R(t_{\text{dec.}})} = 2c\; t_{\text{dec.}}\frac{T_{\text{dec.}}}{T_0} = 2ct_{\text{dec.}} \times 1104 \approx 0.03 \times 2ct_0 \qquad (12.9.1)$$

where $T_{\text{dec.}}$ and T_0 stand respectively for the decoupling temperature of photons and today's temperature of the Universe, and Eq. (12.8.15) has been used in the last step. But the actual CMB region has the maximal diameter $\sim 2ct_0$ at the present time, considerably larger than the expected diameter of the causally connected CMB region $2 \times 1104ct_{\text{dec.}} \approx 0.03 \times 2ct_0$, or far beyond the horizon. In this case, why is the CMB temperature distribution in regions that appear to be causally disconnected so uniform?

12.9.1.2. *The flatness problem*

The total energy density of the Universe Ω_{tot} is observed to be $\sim$0.96, very close to 1 (i.e. $\Omega_{\text{tot}} \sim 1$). In other words, today's Universe is very close to a flat Universe. The Friedmann equation in Eq. (12.1.25) tells us that

$$\Omega_{\text{tot}} - 1 = \frac{\kappa}{\dot{R}^2} \sim \begin{cases} \kappa t^{2/3} \sim \kappa R & \text{(the epoch dominated by matter) } R \sim t^{2/3} \\ \kappa t \sim \kappa R^2 & \text{(the epoch dominated by radiation) } R \sim t^{1/2} \end{cases} \qquad (12.9.2)$$

Since the matter dominance starts from the moment of the matter-radiation equality, $t_{\gamma\text{m}} \approx 6.9 \times 10^4$ yr as given by Eq. (12.8.10), Eq. (12.9.2) leads us to

$$\frac{\Omega_{\text{tot}}(t_{\text{rm}}) - 1}{\Omega_{\text{tot}}(t_0) - 1} = \frac{R(t_{\text{rm}})}{R(t_0)} = \left(\frac{t_{\text{rm}}}{t_0}\right)^{2/3} = \left(\frac{6.9 \times 10^4\text{yr}}{1.4 \times 10^{10}\text{yr}}\right)^{2/3} \approx 3 \times 10^{-4} \qquad (12.9.3)$$

Substituting $\Omega_{\text{tot}}(t_0) - 1 \sim 0.04$ into Eq. (12.9.3), we get

$$\Omega_{\text{tot}}(t_{\text{rm}}) - 1 \sim 1 \times 10^{-5} \qquad (12.9.4)$$

implying that $\Omega_{\text{tot}} - 1 < 10^{-5}$ holds about 6.9×10^4 years after the Big Bang (i.e. Ω_{tot} is closer to 1 at this time).

Given the much earlier Planck time $t_{\text{Pl}} \approx 10^{-43}$ s during the radiation-dominated epoch of the Universe, one may use Eqs. (12.9.2) and (12.9.3) to obtain

$$\begin{aligned}\frac{\Omega_{\text{tot}}(t_{\text{Pl}}) - 1}{\Omega_{\text{tot}}(t_0) - 1} &= \frac{\Omega_{\text{tot}}(t_{\text{Pl}}) - 1}{\Omega_{\text{tot}}(t_{\text{rm}}) - 1} \cdot \frac{\Omega_{\text{tot}}(t_{\text{rm}}) - 1}{\Omega_{\text{tot}}(t_0) - 1} \\ &= \frac{t_{\text{Pl}}}{t_{\text{rm}}} \cdot 3 \times 10^{-4} \\ &\approx \frac{10^{-43}\,\text{s}}{6.9 \times 10^4 \times 3.16 \times 10^7\,\text{s}} \times 3 \times 10^{-4} \\ &\approx 1.4 \times 10^{-59} \qquad (12.9.5)\end{aligned}$$

This result indicates that the fine-tuning is required to be of the 10^{-59} precision level at the Planck time so as to achieve $\Omega_{\text{tot}}(t_0) \sim 1$ today. Such a mysterious fine-tuning is very unnatural from a theoretical point of view, and hence it constitutes the flatness problem in Big Bang cosmology.

12.9.1.3. *The magnetic monopole problem*

The concept of magnetic monopoles was first put forward by Dirac in 1931 [38]. But the Dirac magnetic monopole is a U(1) magnetic monopole with the singularity string. Some GUTs have predicted the existence of magnetic monopoles without singularity, such as an SU(2) magnetic monopole. A magnetic monopole is expected to be stable and have mass $\sim 10^{16} - 10^{17}$ GeV in the GUTs, but such a magnetic monopole has never been observed in any experiments. Why is that?

12.9.2. *The inflation mechanism*

Although the standard Big Bang model has successfully explained many observational data, such as the synthesis of light elements, the existence of the CMB radiation and the homogeneous and isotropic expansion of the Universe, it does not provide the solutions to the aforementioned three problems. To solve these problems, one needs to assume that the vacuum energy dominates the energy density in the very early Universe, so that the Universe may undergo an exponential expansion in a short period of time which is referred as the period of inflation.

If the vacuum energy plays the dominant role in the every early Universe, the Friedmann equation allows us to derive

$$H = \dot{R}/R = \sqrt{\Lambda/3}, \quad \Lambda = 8\pi G_{\rm N}\rho_{\rm V} \tag{12.9.6}$$

and

$$R(t) = R(t_i){\rm e}^{H(t-t_i)} \tag{12.9.7}$$

where t_i denotes the initial time of inflation. It is obvious that $\ddot{R}(t) > 0$ holds; namely, the expansion of the Universe is accelerating. In fact, the density of vacuum energy $\rho_{\rm V}$ is not necessarily a constant. It can be related to the minimum of the potential $V(\phi)$ of a kind of scalar field $\phi(x)$. Replacing $\Lambda/3$ in Eq. (12.1.15) with ρ_Λ and putting it into the total energy density ρ, one obtains

$$\frac{\ddot{R}}{R} = -\frac{4\pi G_{\rm N}}{3}(\rho + 3p) \tag{12.9.8}$$

If $w < -1/3$ holds in the equation of state $p = w\rho$, then one will be left with $\ddot{R}(t) > 0$ from the above equation, making the accelerating expansion of the Universe possible. The inflation mechanism was first proposed by Alan Guth [39], and it was later improved by Linde [40], Albrecht and Steinhardt [41].

Suppose that the density of vacuum energy $\rho_{\rm V}$ comes from the potential $V(\phi)$ of a scalar field ϕ, whose configuration can be taken as the one shown in Fig. 12.8. In this particular case, the scalar field $\phi(x)$ is referred to as "inflaton". When $\phi(x)$ takes its extremum around $\phi \approx 0$, which is a metastable state (false vacuum), one is left with $V(\phi) > 0$. Then the Universe starts to inflate. Later on $\phi(x)$ "rolls" slowly down to the true

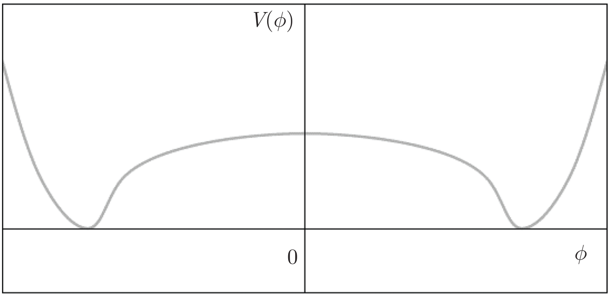


Fig. 12.8. A particular scalar potential with a plateau around $\phi \approx 0$.

vacuum state. When $V(\phi) = 0$, inflation will end. Such an inflation mechanism can help solve the aforementioned three problems. For the sake of convenience, let us first discuss how to deal with the flatness problem.

12.9.2.1. *A solution to the flatness problem*

Suppose that inflation takes place from time t_i to time t_f in the very early Universe (see the plateau of $V(\phi)$ in Fig. 12.8). Then Eqs. (12.9.2) and (12.9.7) allow us to obtain

$$\begin{aligned}\frac{\Omega_{\rm tot}(t_0) - 1}{\Omega_{\rm tot}(t_i) - 1} &= \frac{\Omega_{\rm tot}(t_0) - 1}{\Omega_{\rm tot}(t_{\rm rm}) - 1} \cdot \frac{\Omega_{\rm tot}(t_{\rm rm}) - 1}{\Omega_{\rm tot}(t_f) - 1} \cdot \frac{\Omega_{\rm tot}(t_f) - 1}{\Omega_{\rm tot}(t_i) - 1} \\ &= \left(\frac{t_0}{t_{\rm rm}}\right)^{2/3} \cdot \left(\frac{t_{\rm rm}}{t_f}\right) \cdot {\rm e}^{-2H(t_f - t_i)} \qquad (12.9.9)\end{aligned}$$

and

$$|\Omega_{\rm tot}(t_0) - 1| = |\Omega_{\rm tot}(t_i) - 1| \, {\rm e}^{-2H(t_f - t_i)} \left(\frac{t_{\rm rm}}{t_f}\right) \left(\frac{t_0}{t_{\rm rm}}\right)^{2/3} \qquad (12.9.10)$$

The values of t_f and t_i can be taken around the GUT scale $T \sim 10^{16}$ GeV; namely, $t_i \sim 10^{-38}$s and $t_f \sim 10^{-36}$s. Before t_i the Universe is assumed to be dominated by radiation, and thus $H = \frac{1}{2t}$ holds. So $H(t_i) \sim \frac{1}{t_i}$ at the beginning of inflation, and then the exponential factor in Eq. (12.9.10) is found to be ${\rm e}^{-2H(t_f - t_i)} \sim {\rm e}^{-2\frac{(t_f - t_i)}{t_i}} = {\rm e}^{-200}$. Substituting this result into Eq. (12.9.10), we arrive at

$$|\Omega_{\rm tot}(t_0) - 1| \approx |\Omega_{\rm tot}(t_i) - 1| \times 10^{-35} \qquad (12.9.11)$$

It is clear that $\Omega_{\rm tot}(t_0) \sim 1$ today can be naturally obtained even though $|\Omega_{\rm tot}(t_i) - 1|$ were of $O(1)$ in the very early Universe, so there is no need of making any fine-tuning of the value of $\Omega_{\rm tot}(t_i)$.

12.9.2.2. *A solution to the horizon problem*

Let us assume inflation to take place during $t_i \sim 10^{-38}$s $\rightarrow t_f \sim 10^{-36}$s again. Given $H \sim \frac{1}{t_i}$, the horizon distance at t_i is found to be

$$d_{\rm H}(t_i) = 2c \, t_i \sim 6 \times 10^{-30}{\rm m} \qquad (12.9.12)$$

where c is the speed of light. This is the maximal distance between any two causally connected events at t_i.

During inflation a region of size $d_{\mathrm{H}}(t_i)$ will be enlarged by a factor

$$\frac{R(t_f)}{R(t_i)} = \mathrm{e}^{H(t_f - t_i)} \approx \mathrm{e}^{\frac{t_f - t_i}{t_i}} \sim \mathrm{e}^{100} = 2.7 \times 10^{43} \tag{12.9.13}$$

After inflation the Universe first expands from the radiation-dominated epoch to the matter-radiation equality $t_{\mathrm{rm}}(R \sim t^{1/2},\, t_f \to t_{\mathrm{rm}})$, and then expands to the matter-dominated epoch until today $(R \sim t^{2/3},\, t_{\mathrm{rm}} \to t_0)$. The causally connected region in the beginning of the Universe first inflates and then expands until today, so its size becomes

$$d(t_0) = d_{\mathrm{H}}(t_i) \cdot \mathrm{e}^{H(t_f - t_i)} \left(\frac{t_{\mathrm{rm}}}{t_f}\right)^{1/2} \left(\frac{t_0}{t_{\mathrm{rm}}}\right)^{2/3} \sim 6 \times 10^{41}\,\mathrm{m} \tag{12.9.14}$$

The current Hubble distance $d_{\mathrm{H}}(t_0) = ct_0 \sim 1.3 \times 10^{26}$m is much smaller than the value of $d(t_0)$ given above, and thus the surface of last scattering of the CMB can easily fit into a much larger region characterized by $d(t_0)$. That is why, the CMB radiation is so homogeneous and isotropic in today's Universe.

12.9.2.3. *A solution to the magnetic monopole problem*

Suppose that magnetic monopoles are produced before or during the period of inflation. Around the GUT scale $T \sim 10^{16}$ GeV, inflation is assumed to take place from $t_i \sim 10^{-38}$s to $t_f \sim 10^{-36}$s. Then the number densities of magnetic monopoles at t_f and t_i satisfy the relation

$$n(t_f) = n(t_i)\mathrm{e}^{-3H(t_f - t_i)} \tag{12.9.15}$$

where $n(t) \propto R^{-3}(t)$ has been used. Taking $H \sim \frac{1}{t_i}$, we arrive at

$$n(t_f) \approx \mathrm{e}^{-300} n(t_i) \approx 5 \times 10^{-131} n(t_i) \tag{12.9.16}$$

So the present-day number density of magnetic monopoles is vanishingly small, making it impossible to find out even a single magnetic monopole in the observable Universe.

The above inflation mechanism tells us that the energy density during the period of inflation is $V(\phi) \approx$ constant > 0 in any region of the Universe. After inflation this kind of energy is transferred to particles such as photons, electrons and neutrinos, but their energy densities turn to decrease in a manner of $\rho \sim R^{-4}$ as the Universe continues to expand in a relatively mild way. But inflation itself might not simultaneously begin and end everywhere, leading to spatial variations or fluctuations of the energy density in different regions of the Universe. Such primordial density fluctuations are later on amplified by gravity, and they finally result in the formation of galaxies, galactic clusters and superclusters. They also give rise to the observed CMB anisotropies at the 10^{-5} level in the microwave sky.

This chapter is not intended to cover all the topics in cosmology. Instead, we focus on some important issues which are closely related to particle physics. A comprehensive review of some other interesting issues, such as some hot topics in neutrino astronomy, can be found elsewhere (see, e.g., [7,16–18]).

References

1. PDG Group. *Rev. Part. Phys, Chinese Phys.*, 2014, C38, (9): 090001.
2. Abbot B P, *et al.* (LIGO Scientific Collab. and Vigo Collab.) *Phys. Rev. Lett.*, 2016, 116: 061102; *Astrophys.* J, 2016, 818(2): L22.
3. Hubble E. *Proc. Natl. Acad. Sci.*, 1929, 15: 168.
4. Gamow G, *Phys. Rev*, 1946, 70: 572; Alpher R A. *Phys. Rev*, 1948, 73: 803.
5. Alpher R A, Herman R C. *Phys. Rev*, 1948, 74: 1737; Alpher R A, Herman R C. *Phys. Rev.*, 1949, 75: 1089.
6. Ade P A R, *et al.* Planck Collab. 2013 XVI, arXiv:1303.5076v1; Lahav O, Liddle A R. *The Cosmological Parameters* in PDG Group. *Rev. Part. Phys, Chinese Phys.*, 2014, C38, (9): 090001.
7. Weinberg S. *Gravitation and Cosmology.* John Wiley & Sons. 1972.
8. Weinberg S. *Cosmology.* Oxford University Press, 2008, Errata: http://zippy.ph.utexas.edu/~weinberg/corrections.html/.
9. Friedmann A. Z. *Phys.*, 1992, 10: 377.
10. Young B L. Lecture Notes, *A Survey of Dark Matter and Related Topics in Cosmology*, 2016. (unpublished)
11. Linder E V. *First Principles of Cosmology.* Addison-Wesley Press, 1997: 22–23.
12. Ade P A R, *et al.* Planck Collab. 2013 XVI, arXiv: 1303.5076v1, 2013.
13. Mather J C, *et al. Astrophys. J.*, 1999, 512: 511.
14. Penzias A A, Wilson R W. *Astrophys. J.*, 1965, 142: 419.
15. http://ned.ipac.caltech.edu/level5/Sept05/Gawiser2.html#Figure 1.
16. Xing Z Z, Zhou S. *Neutrinos in Particle Physics, Astronomy and Cosmology.* Springer-Verlag and Zhejiang University Press, 2011.
17. Perkins D H. *Introduction to High Energy Physics.* 4th edition. Cambridge Univ. Press, 2000.
18. Grupen C, *et al. Astroparticle Phys.* Springer-Verlag, 2005.
19. Olive C A, Skillmann E. *Astrophys. J.*, 2004, 617: 29.
20. Permbert M, *et al. Astrophys. J.*, 2007, 667: 636.
21. Izotov Y I, *et al. Astrophys. J.*, 2007, 662: 15.
22. Aver E, *et al. JCAP*, 2012, 04: 004.
23. Aver E, *et al. JCAP*, 2013, 11: 017.

24. Cirelli M, *et al. Nucl. Phys. B*, 2005, 708: 215; Strumia A, Vissani F. arXiv: hep-ph/0606054, 2006.
25. Komatsu E, *et al. Astrophys. J. Suppl.*, 2009, 180: 330.
26. Sakharov A D, *JETP Lett.*, 1967, 5: 24.
27. Alcock C, *et al. Nature*, 1993, V365: 621.
28. Riess A G, *et al. Astrophys. J.*, 1998, 116: 1009; Perlmutter S, *et al. Astrophys. J.*, 1999, 517: 565; Riess A G. *Pub. Astron. Soc. Pac.*, 2000, 112: 1284; Riess A G, *et al. Astrophys. J.*, 2007, 659: 98; Davis T M, *et al. Astrophys. J.*, 2007, 666: 716.
29. Weinberg S. *Rev. Mod. Phys.*, 1989, 60, 1.
30. Ratra B, Peebles P J E. *Rev. Mod. Phys.*, 2003, 75: 559.
31. Smoot G F, *et al. Astrophys. J.*, 1992, 396: L1.
32. Perkins D H. *Particle Astrophysics.* Oxford University Press, 2009; Saha M N. *Proc. Roy. Soc. Lond.*, 1921, A99: 135.
33. Bennett C L, *et al.* WMAP Collab. *Astrophys. J. Supp.*, 2003, 148: 1; Jarosik N, *et al.* WMAP Collab. *Astrophys. J. Supp.*, 2007, 170: 263; Hinshaw G, *et al.* WMAP Collab. *Astrophys. J. Supp.*, 2009, 180: 225.
34. Mather J C, *et al. Astrophys. J.*, 1999, 512: 511.
35. Hinshaw, *et al.* WMAP Collab. *Astrophys. J. Supp.*, 2009, 180: 225.
36. Grupen C, *et al. Astroparticle Phys.* Springer-Verlag, 2005. Fig. 11.4 in P239.
37. Sachs R K, Wolfe A M. *Astrophys. J.*, 1967, 147: 73.
38. Dirac P A M. *Proc. Roy. Soc A.*, 1931, 133: 60.
39. Guth A H. *Phys. Rev D.*, 1981, 23: 347.
40. Linde A D. *Phys. Lett. B*, 1982, 108: 389.
41. Albrecht A, Steinhardt P J. *Phys. Rev. Lett.*, 1982 48: 1220.
42. EHT Collaboration. *APJL*, 2019, 875: L1; L2 (2019a); L3 (2019b); L4 (2019c); L5 (2019d); L6 (2019e).

This page intentionally left blank

Chapter 13

Epilogue

How small is the microscopic world? The electron–positron scattering experiments at the Large Electron–Positron Collider at the Center for European Nuclear Research (CERN) showed that no internal structure of electron was found at the scale of 10^{-17}cm, i.e. the electron is still being thought of as a point-like particle. The high-energy electron at the German e–p collider (HERA) can probe the internal structure of proton deep to 10^{-18} cm. We still do not see the "bottom" of the microscopic world (we do not see next deep layer of matter such as sub-quark or preon). We find from the internal structure research started from atoms $\to$ nuclei $\to$ "elementary particles" $\to$ quarks and leptons $\to \cdots$ that the form of matter deep to every layer does not repeat "the sun system-like structure". On the contrary, there is new feature at every new layer. For example, the nuclei bounded from proton and neutron by strong interactions are more like a pot of porridge, not like the atoms where electrons and nucleus have very loose structure. The quark–gluon structure of elementary particles like the proton and neutron is another new form: quarks cannot be kicked out for being a free moving particles because the confinement mechanism. This makes the internal structure of hadrons more mysterious, so it produces a new form of matter — quark–gluon plasma at heavy ion collision (at RHIC). Up to now, we have not probed the internal structure of quarks and leptons such as electron, μ, τ and neutrinos. How does the internal structure of quarks and leptons look like? It may be fond out at the future experiments of super-high cosmic ray or very high energy e^+e^- linear collider. For particle physics which investigates microscopic structure of matter many countries in the world invest a lot of money. Japan is constructing “supper B factory”, CERN is running the large hadron collider (LHC) and discussing the possibility of constructing Higgs Factory in the near future. China is also discussing the possibility of constructing e^+e^- collider (Higgs Factory) at the center-of-mass energy 250 GeV which can be upgraded to 100 TeV p–p collider SppC (Super pp Collider). The European and American countries are also investigating the possibility of constructing μ storage ring(Neutrino Factory) and the round ring $\mu^+\mu^-$ collider (Higgs Factory). The advantages of the $\mu^+\mu^-$ collider (Higgs Factory) compared with e^+e^- collider (Higgs Factory) are that the energy loss and the same energy machine radius of the former are much smaller than the latter. The key problem is how to obtain the collimated μ beam (beam cooling problem). This problem has been not solved, it is under study. The above plans are pushing the research of the microstructure of matter to deeper and deeper layers.

How large is our macroscopic world? Does our Universe have a “boundary”?

Today our observation to the Universe already extends to the distance of 13 billion light year (one light year is the distance for light to travel one year, i.e. 10,000 billion kilometers), we still do not see the “boundary” of the Universe! Many countries in the world (including China) build or are building super-large astronomic telescopes or satellites for exploring the remote area of the Universe.

From the above discussions, we see that we have not see the boundaries both in deep smaller microscopic direction or in large-scale macroscopic direction. The researches extended to both directions will continue!

When we explore our Universe into both small (microscopic) and large (macroscopic) directions, a subtle phenomenon happens: the extensions to both directions gradually unify. The science exploring “small” direction — particle physics is applying to the science exploring “large” direction — cosmology and astrophysics. The inflation of the Universe after the Big Bang, the formation and development of matter and galaxies of the Universe, the asymmetry of matter and anti-matter (baryon and antibaryon) etc., all of them are connected closely with particle physics. We can say that particle physics makes “small” universe and “large” universe connected. This phenomenon forms many new crossing disciplines of particle physics and astrophysics, etc. These new directions have bright future! Particle physics has also bright future!

In June of 2014, American particle physics Prioritization Panel (P5) suggested five research directions for next ten years:

(1) Use the Higgs boson as a new tool for discovery;
(2) Pursue the physics associated with neutrino mass;
(3) Identify the new physics of dark matter;
(4) Understand cosmic acceleration: dark energy and inflation;
(5) Explore the unknown: new particles, interactions, and physical principles

These priority directions are worthy of reference and investigation.

Finally, we present unsolved ten problems as the ending of the book.

Unsolved Ten Problems

(1) Can we unify all interactions in nature?
(2) Does there exist extra-dimension space–time?
(3) Why do exist so many different kinds of particles?
(4) Why are there only three generations of quarks and leptons?
(5) What is the origin of particle masses? The discovery of the Higgs particle does not solve this problem completely.
(6) What is dark matter? How do we find it?
(7) What is dark energy? If it is some scalar field, how to find it in experiments?
(8) What is the role played by neutrino in the development of the Universe?
(9) Why is there only matter and no antimatter in the Universe today?
(10) Are there new interactions and new physical laws?

To answer these questions, we need common efforts of particle physicists and astrophysicists. We still have long long way to go!

This page intentionally left blank

Appendix I

Calculation of Some Cross Sections

I.1. Calculation of the Cross Section of Electron–Proton Elastic Scattering with Proton Being Viewed as a Point Particle

Calculating the modular square of the amplitude in Eq. (6.4.1), and summing up all the possible spin state for the particles, we can obtain

$$\begin{aligned}\sum_{\alpha\beta\rho\sigma}|A|^2 &= \sum_{\alpha\beta\rho\sigma}\frac{l^4}{q^4}[\bar{u}^\alpha(p)\gamma^\mu u^\beta(p')\bar{u}^\rho(k)\gamma_\mu u^\sigma(k')]\\ &\quad\times[\bar{u}^\beta(p')\gamma^\nu u^\alpha(p)\bar{u}^\sigma(k')\gamma_\nu u^\rho(k)]\\ &= \sum_{\alpha\beta\rho\sigma}\frac{l^4}{q^4}[\bar{u}^\alpha(p)\gamma^\mu u^\beta(p')\bar{u}^\beta(p')\gamma^\nu u^\alpha(p)]\\ &\quad\times[\bar{u}^\rho(k)\gamma_\mu u^\sigma(k')\bar{u}^\sigma(k')\gamma_\nu u^\rho(k)]\\ &= \sum_{\alpha\beta\rho\sigma}\frac{l^4}{q^4}\mathrm{Tr}[u^\alpha(p)\bar{u}^\alpha(p)\gamma^\mu u^\beta(p')\bar{u}^\beta(p')\gamma^\nu]\\ &\quad\times\mathrm{Tr}[u^\rho(k)\bar{u}^\rho(k)\gamma_\mu u^\sigma(k')\bar{u}^\sigma(k')\gamma_\nu]\\ &= \frac{l^4}{q^4}\mathrm{Tr}\left[\frac{\not{p}+M_p}{2M_p}\gamma^\mu\frac{\not{p}'+M_p}{2M_p}\gamma^\nu\right]\mathrm{Tr}[\not{k}\gamma_\mu\ \not{k}'\gamma_\nu]\end{aligned}\tag{I.1.1}$$

where α, β, ρ and σ are the spin indices of the particles in the initial and final states, and the relations satisfied by Dirac spinors are used in the last step of the above equation.

For fermion with mass M_p, there is

$$\sum_{\alpha} u^{\alpha}(p)\bar{u}^{\alpha}(p) = \frac{\not{p} + M_p}{2M_p} \tag{I.1.2}$$

For massless fermion, there is

$$\sum_{\rho} u^{\rho}(k)\bar{u}^{\rho}(k) = \not{k} \tag{I.1.3}$$

Finishing the trace calculation in Eq. (I.1.1), and simplifying it, we can get

$$\sum_{\alpha\beta\rho\sigma} |A|^2 = \frac{e^4}{q^4}\frac{8}{M_p^2}(p \cdot kp' \cdot k' + p \cdot k'p' \cdot k - M_p^2 k \cdot k') \tag{I.1.4}$$

The scattering of electron–proton is equivalent to the scattering between a massless and massive fermions in the case of ignoring the mass of electron. For such a scattering, a change of Eq. (3.4.14) needs to be made in Eq. (3.4.13), i.e. the factor $\frac{1}{4}$ should be changed to $\frac{M_p}{2}$. Then summing up the spin indices of the initial states, we can obtain

$$\mathrm{d}\sigma = \frac{1}{4}\frac{M_p}{2p \cdot k}\sum_{\alpha\beta\rho\sigma} |A|^2 \frac{\mathrm{d}^3p'}{(2\pi)^3}\frac{M_p}{E'_p}\frac{\mathrm{d}^3k'}{(2\pi)^3}\frac{1}{2E'_k}(2\pi)^4\delta^4(k + p - k' - p') \tag{I.1.5}$$

where the factor $\frac{1}{4}$ is for the average of the spin states of the initial particles, and the electron mass has been dropped here, i.e. $m_{\mathrm{e}} \to 0$. Perform the integration for the phase space of the particles in the final state, then we can obtain the total cross section

$$\sigma = \frac{1}{4}\iint \frac{M_p}{2p \cdot k}\sum_{\alpha\beta\rho\sigma} |A|^2 \frac{\mathrm{d}^3p'}{(2\pi)^3}\frac{M_p}{E'_p}\frac{\mathrm{d}^3k'}{(2\pi)^3}\frac{1}{2E'_k}(2\pi)^4\delta^4(k + p - k' - p') \tag{I.1.6}$$

Substituting Eq. (I.1.4) into Eq. (I.1.6), and using

$$\frac{1}{2E'_p} = \int \mathrm{d}p'^0 \delta(p'^2 - M_p^2)\theta(p'^0) \tag{I.1.7}$$

to change $\frac{\mathrm{d}^3p'}{(2\pi)^3}\frac{1}{E'_p}$ in Eq. (I.1.6) to $2\frac{\mathrm{d}^4p'}{(2\pi)^3}\delta(p'^2 - M_p^2)\theta(p'^0)$, then integrate d^4p', we get

$$\begin{aligned}\sigma = \int \frac{e^4}{q^4}\frac{1}{p\cdot k}(p\cdot kp'\cdot k' + p\cdot k'p'\cdot k \\ - M_p^2 k\cdot k')\frac{\mathrm{d}^3k'}{(2\pi)^2}\frac{1}{E'_k}\delta((k+p-k')^2 - M_p^2)\end{aligned} \tag{I.1.8}$$

where $p' = k + p - k'$. Write the integral measure for the electron in the final state d^3k' as $k'^2\mathrm{d}k'\mathrm{d}\Omega$, where $k' = |\vec{k'}|$, and Ω is the solid angle in the momentum space of the final state electron. In the case of dropping the mass of the electron, we have

$$\frac{\mathrm{d}^3k'}{(2\pi)^2}\frac{1}{E'_k} = \frac{k'^2\mathrm{d}k'\mathrm{d}\Omega}{(2\pi)^2}\frac{1}{E'_k} = \frac{E'_k\mathrm{d}E'_k\mathrm{d}\Omega}{(2\pi)^2} \tag{I.1.9}$$

So

$$\begin{aligned}\sigma = \int \frac{e^4}{q^4}\frac{1}{p\cdot k}(p\cdot kp'\cdot k' + p\cdot k'p'\cdot k - M_p^2 k\cdot k')\frac{E'_k\mathrm{d}E'_k\mathrm{d}\Omega}{(2\pi)^2} \\ \times\,\delta((k+p-k')^2 - M_p^2)\end{aligned} \tag{I.1.10}$$

where the δ function can be expanded further

$$\delta((k+p-k')^2 - M_p^2) = \delta(2k\cdot p - 2k\cdot k' - 2p\cdot k') \tag{I.1.11}$$

In laboratory coordinate frame, the proton is at rest, so we have

$$\begin{cases} k\cdot p = M_pE_k \\ k\cdot k' = E_kE'_k(1-\cos\theta) = 2E_kE'_k\sin^2\dfrac{\theta}{2} \\ p\cdot k' = M_pE'_k \end{cases} \tag{I.1.12}$$

where θ is the scattering angle of the electron, i.e. the angle between the momenta of electrons in the final and initial states, and E_k and E'_k are the energies of the initial and final state electrons, respectively.

With Eq. (I.1.12), Eq. (I.1.11) becomes

$$\begin{aligned}\delta((k+p-k')^2-M_p^2) &= \delta\left(2M_pE_k - 4E_k\cdot E_k'\sin^2\frac{\theta}{2} - 2M_p\cdot E_k'\right)\\ &= \delta\left(2M_pE_k - \left(2M_p + 4E_k\sin^2\frac{\theta}{2}\right)E_k'\right)\\ &= \frac{1}{2M_p+4E_k\sin^2\frac{\theta}{2}}\delta\left(\frac{2M_pE_k}{2M_p+4E_k\sin^2\frac{\theta}{2}} - E_k'\right)\\ &= \frac{1}{2\left(M_p+2E_k\sin^2\frac{\theta}{2}\right)}\delta\left(\frac{E_k}{1+\frac{2E_k}{M_p}\sin^2\frac{\theta}{2}} - E_k'\right)\end{aligned} \tag{I.1.13}$$

Substituting Eq. (I.1.13) into Eq. (I.1.10) results in

$$\begin{aligned}\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \int \frac{e^4}{q^4}\frac{1}{p\cdot k}(p\cdot kp'\cdot k' + p\cdot k'p'\cdot k - M_p^2k\cdot k')\frac{E_k'\mathrm{d}E_k'}{(2\pi)^2}\\ \cdot\frac{1}{2\left(M_p+2E_k\sin^2\dfrac{\theta}{2}\right)}\delta\left(\frac{E_k}{1+\frac{2E_k}{M_p}\sin^2\frac{\theta}{2}} - E_k'\right)\end{aligned} \tag{I.1.14}$$

Integrating the energy of the electron, we get

$$\begin{aligned}\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{e^4}{8\pi^2}\frac{1}{q^4p\cdot k}(p\cdot kp'\cdot k' + p\cdot k'p'\cdot k - M_p^2k\cdot k')\frac{E_k}{1+\frac{2E_k}{M_p}\sin^2\frac{\theta}{2}}\\ \cdot\frac{1}{2\left(M_p+2E_k\sin^2\frac{\theta}{2}\right)}\end{aligned} \tag{I.1.15}$$

It can be seen from the δ function in Eq. (I.1.13) that the energy of the final state electron is related to the initial state electron energy and the scattering angle

$$E_k' = \frac{E_k}{1+\frac{2E_k}{M_p}\sin^2\frac{\theta}{2}} \tag{I.1.16}$$

Next let us figure out $p \cdot k$, $p' \cdot k'$, $p \cdot k'$, $p' \cdot k$ and $k \cdot k'$ in Eq. (I.1.15)

$$p \cdot k = M_p E_k \tag{I.1.17a}$$

$$p \cdot k' = M_p E_k' = \frac{M_p E_k}{1 + \frac{2E_k}{M_p} \sin^2 \frac{\theta}{2}} \tag{I.1.17b}$$

$$k \cdot k' = E_k E_k'(1 - \cos\theta) = 2E_k E_k' \sin^2 \frac{\theta}{2} = \frac{2E_k^2 \sin^2 \frac{\theta}{2}}{1 + \frac{2E_k}{M_p} \sin^2 \frac{\theta}{2}} \tag{I.1.17c}$$

$$\begin{aligned} p' \cdot k &= (p + k - k') \cdot k = p \cdot k - k \cdot k' = M_p E_k - 2E_k E_k' \sin^2 \frac{\theta}{2} \\ &= M_p E_k - \frac{2E_k^2 \sin^2 \frac{\theta}{2}}{1 + \frac{2E_k}{M_p} \sin^2 \theta} \end{aligned} \tag{I.1.17d}$$

$$\begin{aligned} p' \cdot k' &= (p + k - k') \cdot k' = p \cdot k' + k \cdot k' \\ &= \frac{M_p E_k}{1 + \frac{2E_k}{M_p} \sin^2 \frac{\theta}{2}} + \frac{2E_k^2 \sin^2 \frac{\theta}{2}}{1 + \frac{2E_k}{M_p} \sin^2 \frac{\theta}{2}} = M_p E_k \end{aligned} \tag{I.1.17e}$$

Substituting Eqs. (I.1.17a–I.1.17e) into Eq. (I.1.15) and simplifying it, we get

$$\begin{aligned} \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{e^4}{8\pi^2} \frac{1}{q^4 M_p^2} &\left(\frac{2(M_p E_k)^2 \cos^2 \frac{\theta}{2}}{1 + \frac{2E_k}{M_p} \sin^2 \frac{\theta}{2}} + \frac{4E_k^4 \sin^4 \frac{\theta}{2}}{\left(1 + \frac{2E_k}{M_p} \sin^2 \frac{\theta}{2}\right)^2} \right) \\ &\cdot \frac{1}{\left(1 + \frac{2E_k}{M_p} \sin^2 \frac{\theta}{2}\right)^2} \end{aligned} \tag{I.1.18}$$

Considering

$$q^2 = (k - k')^2 = -2k \cdot k' = \frac{-4E_k^2 \sin^2 \frac{\theta}{2}}{1 + \frac{2E_k}{M_p} \sin^2 \frac{\theta}{2}} \tag{I.1.19}$$

Eq. (I.1.18) can be expressed in a more concise form

$$\begin{aligned}
\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} &= \frac{e^4}{8\pi^2}\frac{1}{q^4 M_p^2}\left(\frac{1}{2}M_p^2|q^2|\tan^{-2}\frac{\theta}{2}+\frac{1}{4}q^4\right)\frac{1}{\left(1+\frac{2E_k}{M_p}\sin^2\frac{\theta}{2}\right)^2} \\
&= \frac{e^4}{16\pi^2}\frac{1}{|q^2|}\tan^{-2}\frac{\theta}{2}\left(1+\frac{1}{2}\frac{|q^2|}{M_p^2}\tan^2\frac{\theta}{2}\right)\frac{1}{\left(1+\frac{2E_k}{M_p}\sin^2\frac{\theta}{2}\right)^2} \\
&= \alpha^2\frac{1+\frac{2E_k}{M_p}\sin^2\frac{\theta}{2}}{4E_k^2\sin^2\frac{\theta}{2}}\frac{\cos^2\frac{\theta}{2}}{\sin^2\frac{\theta}{2}}\left(1+\frac{1}{2}\frac{|q^2|}{M_p^2}\tan^2\frac{\theta}{2}\right)\frac{1}{\left(1+\frac{2E_k}{M_p}\sin^2\frac{\theta}{2}\right)^2} \\
&= \frac{\alpha^2\cos^2\frac{\theta}{2}}{4E_k^2\sin^4\frac{\theta}{2}}\frac{1}{\left(1+\frac{2E_k}{M_p}\sin^2\frac{\theta}{2}\right)}\left(1+\frac{1}{2}\frac{|q^2|}{M_p^2}\tan^2\frac{\theta}{2}\right) \qquad \text{(I.1.20)}
\end{aligned}$$

If changing the symbol of the energy of the final state electron from E_k to E, and defining $Q^2 \equiv -q^2$, then Eq. (I.1.20) can be written as

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{\alpha^2\cos^2\frac{\theta}{2}}{4E^2\sin^4\frac{\theta}{2}}\frac{1}{\left(1+\frac{2E}{M_p}\sin^2\frac{\theta}{2}\right)}\left(1+\frac{1}{2}\frac{Q^2}{M_p^2}\tan^2\frac{\theta}{2}\right) \qquad \text{(I.1.21)}$$

I.2. Calculation of the Elastic Scattering Cross Section of Electron and a Particle with Spin 0 and Charge 1

The interaction vertex between charged particle of spin 0 and electromagnetic field is $\mathrm{i}Qe(p+p')_\mu$. The Feynman diagram is shown in Fig. 6.5. The scattering amplitude is

$$\begin{aligned}
A &= \bar{u}^\alpha(k')(-\mathrm{i})e\gamma_\mu u^\beta(k)\frac{-\mathrm{i}}{q^2}\mathrm{i}e(p+p')^\mu \\
&= -\mathrm{i}e^2\bar{u}^\alpha(k')\gamma_\mu u^\beta(k)\frac{1}{q^2}(p+p')^\mu \qquad \text{(I.2.1)}
\end{aligned}$$

Calculate the modular square of the amplitude and sum up the spin states for the initial and final state electrons

$$\begin{aligned}
\sum_{\alpha\beta}|A|^2 &= \sum_{\alpha\beta}\left[e^2\bar{u}^\alpha(k')\gamma_\mu u^\beta(k)\frac{1}{q^2}(p+p')^\mu\right]^+ e^2\bar{u}^\alpha(k')\gamma_\nu u^\beta(k)\frac{1}{q^2}(p+p')^\nu \\
&= \sum_{\alpha\beta}e^4\bar{u}^\beta(k)\gamma_\mu u^\alpha(k')\bar{u}^\alpha(k')\gamma_\nu u^\beta(k)\frac{1}{q^4}(p+p')^\mu(p+p')^\nu
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{\alpha\beta} e^4 \mathrm{Tr}[u^\beta(k)\bar{u}^\beta(k)\gamma_\mu u^\alpha(k')\bar{u}^\alpha(k')\gamma_\nu]\frac{1}{q^4}(p+p')^\mu(p+p')^\nu \\
&= e^4 \mathrm{Tr}[\not{k}\gamma_\mu \ \not{k}'\gamma_\nu](p+p')^\mu(p+p')^\nu \frac{1}{q^4} \\
&= 4e^4[k_\mu k'_\nu - k\cdot k' g_{\mu\nu} + k_\nu k'_\mu](p+p')^\mu(p+p')^\nu \frac{1}{q^4} \\
&= 4e^4[2k\cdot(p+p')k'\cdot(p+p') - k\cdot k'(p+p')^2]\frac{1}{q^4} \\
&= 4e^4[2k\cdot(2p+k-k')k'\cdot(2p+k-k') - k\cdot k'(2p+k-k')^2]\frac{1}{q^4} \\
&= 4e^4[2(2k\cdot p - k\cdot k')(2p\cdot k' + k\cdot k') \\
&\quad - k\cdot k'(4M_p^2 + 4p\cdot k - 4p\cdot k' - 2k\cdot k')]\frac{1}{q^4} \\
&= 4e^4[8k\cdot pp\cdot k' - 4M_p^2 k\cdot k']\frac{1}{q^4} \\
&= 16e^4\left[2M_p^2 E_k E'_k - 2M_p^2 E_k E'_k \sin^2\frac{\theta}{2}\right]\frac{1}{q^4} \\
&= 32e^4 M_p^2 E_k E'_k \cos^2\frac{\theta}{2}\frac{1}{q^4} \qquad \text{(I.2.2)}
\end{aligned}
$$

Substituting E'_k in Eq. (I.1.16) into the above result, we get

$$
\sum_{\alpha\beta}|A|^2 = 32e^4\frac{M_p^2 E_k^2\cos^2\frac{\theta}{2}}{1+\frac{2E_k}{M_p}\sin^2\frac{\theta}{2}}\frac{1}{q^4} \qquad \text{(I.2.3)}
$$

According to Eqs. (3.4.13) and (3.4.14), the total scattering cross section of an electron and a scalar particle is

$$
\sigma = \frac{1}{2}\iint \frac{1}{4p\cdot k}\sum_{\alpha\beta}|A|^2\frac{\mathrm{d}^3p'}{(2\pi)^3}\frac{1}{2E'_p}\frac{\mathrm{d}^3k'}{(2\pi)^3}\frac{1}{2E'_k}(2\pi)^4\delta^4(k+p-k'-p') \qquad \text{(I.2.4)}
$$

where the electron mass is neglected, and the factor in front of the above equation is the average factor for the initial particle's spin. Using Eqs. (I.1.7), (I.1.9) and (I.1.13), and repeating the calculations as before, we can finally transform the phase space factor in Eq. (I.2.4) to be

$$\begin{aligned}&\frac{\mathrm{d}^3p'}{(2\pi)^3}\frac{1}{2E'_p}\frac{\mathrm{d}^3k'}{(2\pi)^3}\frac{1}{2E'_k}(2\pi)^4\delta^4(k+p-k'-p')\\&\quad\to\frac{1}{2}\frac{E'_k\mathrm{d}E'_k\mathrm{d}\Omega}{(2\pi)^2}\frac{1}{2\left(M_p+2E_k\sin^2\frac{\theta}{2}\right)}\delta\left(\frac{E_k}{1+\frac{2E_k}{M_p}\sin^2\frac{\theta}{2}}-E'_k\right)\end{aligned}\tag{I.2.5}$$

Substituting the result of Eq. (I.2.5) into Eq. (I.2.4) can lead to

$$\begin{aligned}\sigma=\frac{1}{2}\iint&\frac{1}{4p\cdot k}\sum_{\alpha\beta}|A|^2\frac{1}{2}\frac{E'_k\mathrm{d}E'_k\mathrm{d}\Omega}{(2\pi)^2}\frac{1}{2(M_p+2E_k\sin^2\frac{\theta}{2})}\\&\times\delta\left(\frac{E_k}{1+\frac{2E_k}{M_p}\sin^2\frac{\theta}{2}}-E'_k\right)=\int\frac{1}{16p\cdot k}\sum_{\alpha\beta}|A|^2\frac{\mathrm{d}\Omega}{(2\pi)^2}\\&\times\frac{E_k}{1+\frac{2E_k}{M_p}\sin^2\frac{\theta}{2}}\frac{1}{2(M_p+2E_k\sin^2\frac{\theta}{2})}=\int\frac{1}{16M_pE_k}\sum_{\alpha\beta}|A|^2\frac{\mathrm{d}\Omega}{(2\pi)^2}\\&\times\frac{E_k}{1+\frac{2E_k}{M_p}\sin^2\frac{\theta}{2}}\frac{1}{2(M_p+2E_k\sin^2\frac{\theta}{2})}=\int\frac{1}{16M_p}\sum_{\alpha\beta}|A|^2\frac{\mathrm{d}\Omega}{(2\pi)^2}\\&\times\frac{1}{1+\frac{2E_k}{M_p}\sin^2\frac{\theta}{2}}\frac{1}{2(M_p+2E_k\sin^2\frac{\theta}{2})}\end{aligned}\tag{I.2.6}$$

Then the differential cross section can be obtained from the result in the above equation

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}=\frac{1}{64\pi^2M_p}\sum_{\alpha\beta}|A|^2\frac{1}{1+\frac{2E_k}{M_p}\sin^2\frac{\theta}{2}}\frac{1}{2(M_p+2E_k\sin^2\frac{\theta}{2})}\tag{I.2.7}$$

Substituting the result of Eq. (I.2.3) into above, we obtain

$$\begin{aligned}\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}&=\frac{1}{64\pi^2M_p}32e^4\frac{M_p^2E_k^2\cos^2\frac{\theta}{2}}{1+\frac{2E_k}{M_p}\sin^2\frac{\theta}{2}}\frac{1}{q^4}\frac{1}{1+\frac{2E_k}{M_p}\sin^2\frac{\theta}{2}}\frac{1}{2(M_p+2E_k\sin^2\frac{\theta}{2})}\\&=\frac{e^4}{4\pi^2}\frac{E_k^2\cos^2\frac{\theta}{2}}{\left(1+\frac{2E_k}{M_p}\sin^2\frac{\theta}{2}\right)^3}\frac{1}{q^4}\end{aligned}\tag{I.2.8}$$

Considering the result of q^2 in Eq. (I.2.8), one can change Eq. (I.2.8) to

$$\begin{aligned}\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} &= \frac{e^4}{4\pi^2}\frac{E_k^2\cos^2\frac{\theta}{2}}{\left(1+\frac{2E_k}{M_p}\sin^2\frac{\theta}{2}\right)^3}\left(\frac{1+\frac{2E_k}{M_p}\sin^2\frac{\theta}{2}}{4E_k^2\sin^2\frac{\theta}{2}}\right)^2 \\ &= \frac{e^4}{64\pi^2}\frac{\cos^2\frac{\theta}{2}}{E_k^2\sin^4\frac{\theta}{2}}\frac{1}{1+\frac{2E_k}{M_p}\sin^2\frac{\theta}{2}} \\ &= \frac{\alpha^2}{4}\frac{\cos^2\frac{\theta}{2}}{E_k^2\sin^4\frac{\theta}{2}}\frac{1}{1+\frac{2E_k}{M_p}\sin^2\frac{\theta}{2}} \qquad \text{(I.2.9)}\end{aligned}$$

Changing the symbol of the energy of the final state electron from E_k to E, we can obtain

$$\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_{\text{Mot}} = \frac{\alpha^2}{4}\frac{\cos^2\frac{\theta}{2}}{E^2\sin^4\frac{\theta}{2}}\frac{1}{1+\frac{2E}{M_p}\sin^2\frac{\theta}{2}} \qquad \text{(I.2.10)}$$

This result is called Mott scattering formula (N.F. Mott).

I.3. Calculation of the Cross Section of Electron–Proton Elastic Scattering

The cross section of the elastic scattering of electron and a real proton is calculated in this subsection. It is known from Eq. (6.4.10), that the amplitude of the elastic scattering of electron and proton can be written as

$$A = \frac{-\mathrm{i}e^2}{q^2}j^\mu(k',k)J_\mu(p',p) \qquad \text{(I.3.1)}$$

where j^μ and J_μ are the electromagnetic currents for leptons and hadrons, respectively (details can be seen in Eqs. (6.4.8) and (6.4.16))

$$j^\mu(k',k) = \bar{u}^\alpha(k')\gamma^\mu u^\beta(k) \qquad \text{(I.3.2)}$$

$$J_\mu(p',p) = \bar{u}^\rho(p')[\gamma_\mu F_1(q^2) + \mathrm{i}F_2(q^2)\sigma_{\mu\nu}q^\nu]u^\sigma(p) \qquad \text{(I.3.3)}$$

Here α, β, ρ and σ are the spin indices of the initial and final state electron and protons. Calculate the modular square and sum up all the the spin states of the initial and final state particles

$$\sum_{\alpha\beta\rho\sigma}|A| = \sum_{\alpha\beta\rho\sigma}\frac{e^4}{q^4}[j^\mu(k',k)J_\mu(p',p)]^+ j^\nu(k',k)J_\nu(p',p)$$
$$= \sum_{\alpha\beta\rho\sigma}\frac{e^4}{q^4}j^{\mu+}(k',k)j^\nu(k',k)J_\mu^+(p',p)J_\nu(p',p) \qquad \text{(I.3.4)}$$

Define the leptonic and hadronic current tensors as

$$l^{\mu\nu} = \sum_{\alpha\beta} j^{\mu+}(k',k)j^\nu(k',k) \qquad \text{(I.3.5a)}$$

$$h_{\mu\nu} = \sum_{\rho\sigma} J_\mu^+(p',p)J_\nu(p',p) \qquad \text{(I.3.5b)}$$

then Eq. (I.3.4) can be written as

$$\sum_{\alpha\beta\rho\sigma}|A| = \frac{e^4}{q^4}l^{\mu\nu}h_{\mu\nu} \qquad \text{(I.3.6)}$$

Substituting Eq. (I.3.2) into Eq. (I.3.5a), one can calculate the leptonic current tensor

$$l^{\mu\nu} = \sum_{\alpha\beta}\bar{u}^\beta(k)\gamma^\mu u^\alpha(k')\bar{u}^\alpha(k')\gamma^\nu u^\beta(k)$$
$$= \sum_{\alpha\beta}\mathrm{Tr}[u^\beta(k)\bar{u}^\beta(k)\gamma^\mu u^\alpha(k')\bar{u}^\alpha(k')\gamma^\nu]$$
$$= \mathrm{Tr}[\not{k}\gamma^\mu \not{k}'\gamma^\nu]$$
$$= 4(k^\mu k'^\nu - k\cdot k' g^{\mu\nu} + k^\nu k'^\mu)$$
$$= 4\left(k^\mu k'^\nu + k^\nu k'^\mu + \frac{1}{2}q^2 g^{\mu\nu}\right) \qquad \text{(I.3.7)}$$

where $q^2 = (k-k')^2 = -2k\cdot k'$ has been used.

Substituting Eq. (I.3.3) into Eq. (I.3.5b), one can calculate the hadronic current tensor. For this purpose, let us calculate the Hermitian conjugate of the hadronic electromagnetic current in Eq. (I.3.3)

$$\begin{aligned} J_\mu^+(p',p) &= [\bar{u}^\rho(p')[\gamma_\mu F_1(q^2) + \mathrm{i}F_2(q^2)\sigma_{\mu\delta}q^\delta]u^\sigma(p)]^+ \\ &= u^{\sigma+}(p)[\gamma_\mu^+ F_1(q^2) - \mathrm{i}F_2(q^2)\sigma_{\mu\delta}^+ q^\delta]\gamma^0 u^\rho(p') \\ &= \bar{u}^\sigma(p)\gamma^0[\gamma_\mu^+ F_1(q^2) - \mathrm{i}F_2(q^2)\sigma_{\mu\delta}^+ q^\delta]\gamma^0 u^\rho(p') \\ &= \bar{u}^\sigma(p)[\gamma^0\gamma_\mu^+\gamma^0 F_1(q^2) - \mathrm{i}F_2(q^2)\gamma^0\sigma_{\mu\delta}^+\gamma^0 q^\delta]u^\rho(p') \end{aligned} \tag{I.3.8}$$

where the form factors $F_{1,2}(q^2)$ are treated as real numbers. Using

$$\gamma^0\gamma_\mu^+\gamma^0 = \gamma_\mu \tag{I.3.9}$$

and

$$\gamma^0\sigma_{\mu\delta}^+\gamma^0 = \sigma_{\mu\delta} \tag{I.3.10}$$

one can change Eq. (I.3.8) to

$$J_\mu^+(p',p) = \bar{u}^\sigma(p)[\gamma_\mu F_1(q^2) - \mathrm{i}F_2(q^2)\sigma_{\mu\delta}q^\delta]u^\rho(p') \tag{I.3.11}$$

With the above result, we can now calculate the hadronic current tensor. Substituting Eqs. (I.3.11) and (I.3.3) into Eq. (I.3.5b), we get

$$\begin{aligned} h_{\mu\nu} &= \sum_{\rho\sigma} \bar{u}^\sigma(p)[\gamma_\mu F_1(q^2) - \mathrm{i}F_2(q^2)\sigma_{\mu\delta}q^\delta]u^\rho(p')\bar{u}^\rho(p') \\ &\quad \cdot [\gamma_\nu F_1(q^2) + \mathrm{i}F_2(q^2)\sigma_{\nu\gamma}q^\gamma]u^\sigma(p) \\ &= \sum_{\rho\sigma} \mathrm{Tr}[u^\sigma(p)\bar{u}^\sigma(p)[\gamma_\mu F_1(q^2) - \mathrm{i}F_2(q^2)\sigma_{\mu\delta}q^\delta]u^\rho(p')\bar{u}^\rho(p') \\ &\quad \cdot [\gamma_\nu F_1(q^2) + \mathrm{i}F_2(q^2)\sigma_{\nu\gamma}q^\gamma]] \\ &= \mathrm{Tr}\left\{ \frac{\not{p} + M_p}{2M_p}[\gamma_\mu F_1(q^2) - \mathrm{i}F_2(q^2)\sigma_{\mu\delta}q^\delta]\frac{\not{p}' + M_p}{2M_p} \right. \\ &\quad \left. \times [\gamma_\nu F_1(q^2) + \mathrm{i}F_2(q^2)\sigma_{\nu\gamma}q^\gamma] \right\} \\ &= \frac{1}{4M_p^2}\mathrm{Tr}[(\not{p} + M_p)[\gamma_\mu F_1(q^2) - \mathrm{i}F_2(q^2)\sigma_{\mu\delta}q^\delta](\not{p}' + M_p) \\ &\quad \cdot [\gamma_\nu F_1(q^2) + \mathrm{i}F_2(q^2)\sigma_{\nu\gamma}q^\gamma]] \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{4M_p^2}\{\mathrm{Tr}[(\not{p}+M_p)\gamma_\mu(\not{p}'+M_p)\gamma_\nu]F_1^2(q^2) \\
&\quad + \mathrm{Tr}[(\not{p}+M_p)\gamma_\mu(\not{p}'+M_p)\sigma_{\nu\gamma}q^\gamma]\mathrm{i}F_1(q^2)F_2(q^2) \\
&\quad - \mathrm{Tr}[(\not{p}+M_p)\sigma_{\mu\delta}(\not{p}'+M_p)\gamma_\nu q^\delta]\mathrm{i}F_1(q^2)]F_2(q^2) \\
&\quad + \mathrm{Tr}[(\not{p}+M_p)\sigma_{\mu\delta}(\not{p}'+M_p)\sigma_{\nu\gamma}q^\delta q^\gamma]F_2^2(q^2)\} \qquad \text{(I.3.12)}
\end{aligned}$$

where the trace of the first term is

$$\mathrm{Tr}[(\not{p}+M_p)\gamma_\mu(\not{p}'+M_p)\gamma_\nu] = 4(p_\mu p'_\nu + p_\nu p'_\mu - p\cdot p' g_{\mu\nu} + M_p^2 g_{\mu\nu}) \qquad \text{(I.3.13)}$$

Considering $p' = p + q$, $p'^2 = (p+q)^2$ and $p \cdot q = -\frac{1}{2}q^2$, one can finally express the result of Eq. (I.3.13) as

$$\begin{aligned}
&\mathrm{Tr}[(\not{p}+M_p)\gamma_\mu(\not{p}'+M_p)\gamma_\nu] \\
&\quad = 4\left[2\left(p_\mu - \frac{p\cdot q}{q^2}q_\mu\right)\left(p_\nu - \frac{p\cdot q}{q^2}q_\nu\right)\right. \\
&\qquad \left. + \frac{1}{2}(q^2 g_{\mu\nu} - q_\mu q_\nu)\right] \qquad \text{(I.3.14)}
\end{aligned}$$

and the traces of the second, third and fourth terms are

$$\mathrm{Tr}[(\not{p}+M_p)\gamma_\mu(\not{p}'+M_p)\sigma_{\nu\gamma}q^\gamma] = 4\mathrm{i}(-q^2 g_{\mu\nu} + q_\mu q_\nu) \qquad \text{(I.3.15)}$$

$$\mathrm{Tr}[(\not{p}+M_p)\sigma_{\mu\delta}(\not{p}'+M_p)\gamma_\nu q^\delta] = 4\mathrm{i}(q^2 g_{\mu\nu} - q_\mu q_\nu) \qquad \text{(I.3.16)}$$

$$\begin{aligned}
&\mathrm{Tr}[(\not{p}+M_p)\sigma_{\mu\delta}(\not{p}'+M_p)\sigma_{\mu\gamma}q^\delta q^\gamma] \\
&\quad = 4\left[2M_p^2(q^2 g_{\mu\nu} - q_\mu q_\nu) - 2q^2\left(p_\mu - \frac{p\cdot q}{q^2}q_\mu\right)\left(p_\nu - \frac{p\cdot q}{q^2}q_\nu\right)\right] \qquad \text{(I.3.17)}
\end{aligned}$$

With the results of Eqs. (I.3.14)–(I.3.17), Eq. (I.3.12) can be simplified to be

$$\begin{aligned}
h_{\mu\nu} &= \frac{1}{M_p^2}\left[2\left(p_\mu - \frac{p\cdot q}{q^2}q_\mu\right)\left(p_\nu - \frac{p\cdot q}{q^2}q_\nu\right)(F_1^2 - q^2F_2^2)\right. \\
&\qquad \left. + \frac{1}{2}(q^2 g_{\mu\nu} - q_\mu q_\nu)(F_1 + 2M_pF_2)^2\right] \qquad \text{(I.3.18)}
\end{aligned}$$

Substituting Eqs. (I.3.7) and (I.3.18) into Eq. (I.3.6), we can get

$$\sum_{\alpha\beta\rho\sigma}|A| = \frac{4e^4}{M_p^2 q^4}\Bigg[2\left(k^\mu k'^\nu + k^\nu k'^\mu + \frac{1}{2}q^2 g^{\mu\nu}\right)\left(p_\mu - \frac{p\cdot q}{q^2}q_\mu\right)$$

$$\cdot\left(p_\nu - \frac{p\cdot q}{q^2}q_\nu\right)\left(F_1^2 - q^2F_2^2\right) + \frac{1}{2}\left(k^\mu k'^\nu + k^\nu k'^\mu + \frac{1}{2}q^2 g^{\mu\nu}\right)$$

$$\cdot(q^2 g_{\mu\nu} - q_\mu q_\nu)(F_1 + 2M_pF_2)^2\Bigg] \tag{I.3.19}$$

And then we need to deal with the scalar product of the four-momenta, and we are going to use the results of the dot-product of four-momenta. In lab coordinates, proton is stationary, so we have

$$k\cdot p = M_pE_k, \quad p\cdot k' = M_pE'_k, \quad k\cdot q = k\cdot(k-k') = -k\cdot k' = \frac{1}{2}q^2,$$

$$k'\cdot q = k'\cdot(k-k') = k'\cdot k = -\frac{1}{2}q^2, \quad p\cdot q = -\frac{1}{2}q^2,$$

$$q^2 = (k-k')^2 = -2k\cdot k' = -2E_kE'_k(1-\cos\theta) = -4E_kE'_k\sin^2\frac{\theta}{2} \tag{I.3.20}$$

With the above equation, Eq. (I.3.19) can be finally simplified to be

$$\sum_{\alpha\beta\rho\sigma}|A| = \frac{4e^4}{M_p^2 q^4}\Bigg[4M_p^2E_kE'_k\cos^2\frac{\theta}{2}(F_1^2 - q^2F_2^2)$$

$$-2q^2E_kE'_k\sin^2\frac{\theta}{2}(F_1 + 2M_pF_2)^2\Bigg]$$

$$= \frac{16e^4E_kE'_k\cos^2\dfrac{\theta}{2}}{q^4}\left[(F_1^2 - q^2F_2^2) - \frac{q^2}{2M_p^2}(F_1 + 2M_pF_2)^2\tan^2\frac{\theta}{2}\right]$$

$$= \frac{e^4\cos^2\frac{\theta}{2}}{E_kE'_k\sin^4\frac{\theta}{2}}\left[(F_1^2 - q^2F_2^2) - \frac{q^2}{2M_p^2}(F_1 + 2M_pF_2)^2\tan^2\frac{\theta}{2}\right] \tag{I.3.21}$$

Define $Q^2 = -q^2$, then the above equation becomes

$$\sum_{\alpha\beta\rho\sigma}|A| = \frac{e^4\cos^2\frac{\theta}{2}}{E_kE'_k\sin^4\dfrac{\theta}{2}}\left[(F_1^2 + Q^2F_2^2) + \frac{Q^2}{2M_p^2}(F_1 + 2M_pF_2)^2\tan^2\frac{\theta}{2}\right] \tag{I.3.22}$$

Now we can calculate the cross section of electron–proton scattering

$$\sigma = \frac{1}{4}\iint \frac{M_p}{2p\cdot k}\sum_{\alpha\beta\rho\sigma}|A|^2\frac{\mathrm{d}^3p'}{(2\pi)^3}\frac{M_p}{E'_p}\frac{\mathrm{d}^3k'}{(2\pi)^3}\frac{1}{2E'_k}(2\pi)^4\delta^4(k+p-k'-p') \tag{I.3.23}$$

Using

$$\frac{1}{2E'_p} = \int \mathrm{d}p^{'0}\delta(p^{'2}-M_p^2)\theta(p^{'0}) \tag{I.3.24a}$$

and

$$\frac{\mathrm{d}^3k'}{(2\pi)^2}\frac{1}{E'_k} = \frac{k'^2\mathrm{d}k'\mathrm{d}\Omega}{(2\pi)^2}\frac{1}{E'_k} = \frac{E'_k\mathrm{d}E'_k\mathrm{d}\Omega}{(2\pi)^2} \tag{I.3.24b}$$

and performing the integration of $\int \mathrm{d}^4p'$, the phase space factor in the above equation becomes

$$\begin{aligned}&\frac{\mathrm{d}^3p'}{(2\pi)^3}\frac{M_p}{E'_p}\frac{\mathrm{d}^3k'}{(2\pi)^3}\frac{1}{2E'_k}(2\pi)^4\delta^4(k+p-k'-p')\\ &\quad\to \frac{1}{8\pi^2}\frac{E'_k\mathrm{d}E'_k\mathrm{d}\Omega}{1+\frac{2E_k}{M_p}\sin^2\frac{\theta}{2}}\delta\left(\frac{E_k}{1+\frac{2E_k}{M_p}\sin^2\frac{\theta}{2}}-E'_k\right)\end{aligned} \tag{I.3.25}$$

Therefore Eq. (I.3.23) becomes

$$\begin{aligned}\sigma &= \frac{1}{4}\int\frac{M_p}{2p\cdot k}\sum_{\alpha\beta\rho\sigma}|A|^2\frac{1}{8\pi^2}\frac{E'_k\mathrm{d}E'_k\mathrm{d}\Omega}{1+\frac{2E_k}{M_p}\sin^2\frac{\theta}{2}}\delta\left(\frac{E_k}{1+\frac{2E_k}{M_p}\sin^2\frac{\theta}{2}}-E'_k\right)\\ &= \frac{1}{64\pi^2}\int\frac{1}{E_k}\sum_{\alpha\beta\rho\sigma}|A|^2\frac{E'_k\mathrm{d}E'_k\mathrm{d}\Omega}{1+\frac{2E_k}{M_p}\sin^2\frac{\theta}{2}}\delta\left(\frac{E_k}{1+\frac{2E_k}{M_p}\sin^2\frac{\theta}{2}}-E'_k\right)\end{aligned} \tag{I.3.26}$$

Substituting Eq. (I.3.22) into the above equation and simplifying it, we have

$$\begin{aligned}\sigma &= \int \mathrm{d}E'_k\mathrm{d}\Omega\frac{\alpha^2\cos^2\frac{\theta}{2}}{4E_k^2\sin^4\frac{\theta}{2}}\frac{1}{1+\frac{2E_k}{M_p}\sin^2\frac{\theta}{2}}\delta\left(\frac{E_k}{1+\frac{2E_k}{M_p}\sin^2\frac{\theta}{2}}-E'_k\right)\\ &\quad\cdot\left[(F_1^2+Q^2F_2^2)+\frac{Q^2}{2M_p^2}(F_1+2M_pF_2)^2\tan^2\frac{\theta}{2}\right]\end{aligned} \tag{I.3.27}$$

The above result can also be expressed in terms of differential cross section

$$\frac{\mathrm{d}\sigma}{\mathrm{d}E_k'\mathrm{d}\Omega}=\frac{\alpha^2\cos^2\frac{\theta}{2}}{4E_k^2\sin^4\frac{\theta}{2}}\frac{1}{1+\frac{2E_k}{M_p}\sin^2\frac{\theta}{2}}\delta\left(\frac{E_k}{1+\frac{2E_k}{M_p}\sin^2\frac{\theta}{2}}-E_k'\right)\cdot\left[(F_1^2+Q^2F_2^2)+\frac{Q^2}{2M_p^2}(F_1+2M_pF_2)^2\tan^2\frac{\theta}{2}\right]\tag{I.3.28}$$

If we integrate the final electron energy E_k', we get

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}=\frac{\alpha^2\cos^2\frac{\theta}{2}}{4E_k^2\sin^4\frac{\theta}{2}}\frac{1}{1+\frac{2E_k}{M_p}\sin^2\frac{\theta}{2}}\cdot\left[(F_1^2+Q^2F_2^2)+\frac{Q^2}{2M_p^2}(F_1+2M_pF_2)^2\tan^2\frac{\theta}{2}\right]\tag{I.3.29}$$

With the result of Mott scattering formula, the above equation can be written as

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}=\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_{\mathrm{Mot}}\left[(F_1^2+Q^2F_2^2)+\frac{Q^2}{2M_p^2}(F_1+2M_pF_2)^2\tan^2\frac{\theta}{2}\right]\tag{I.3.30}$$

I.4. Calculation of the Cross Section of Eectron–Proton Inelastic Scattering

The diagram for electron–proton inelastic scattering is shown in Fig. 6.6. As shown in Eq. (6.5.2), the amplitude of electron proton inelastic scattering process can be written as

$$A_{\mathrm{NE}}=e^2j_\mu\frac{1}{q^2}J^\mu\tag{I.4.1}$$

where j_μ is the leptonic electromagnetic current

$$j_\mu=\bar{u}(k')\gamma_\mu u(k)\tag{I.4.2}$$

and J^μ is the hadronic electromagnetic current for the inelastic scattering process. To calculate the scattering cross section, the modular square of the scattering amplitude is needed. So we need to calculation the modular square of Eq. (I.4.1), and sum up all the possible spin states for both the initial and final state particles

$$\sum |A_{\mathrm{NE}}|^2 = \sum \frac{e^4}{q^4} j_\mu^+ j_\nu J^{\mu+} J^\nu \tag{I.4.3}$$

where $\sum$ stands for summation over the spin states of the initial and final state particles.

Define the leptonic current tensor averaged for initial states

$$L_{\mu\nu} \equiv \frac{1}{2} \sum_{\lambda_1} \sum_{\lambda_2} j_\mu^+ j_\nu = 2\left(k_\mu k'_\nu + k_\nu k'_\mu + \frac{q^2}{2} g_{\mu\nu}\right) \tag{I.4.4}$$

where $\frac{1}{2}\sum_{\lambda_1}$ is averaging for the spin states of electron of initial state, and $\sum_{\lambda_2}$ summation for the spin states of final state electron. The unpolarized hadronic current tensor is defined as (see also in Eq. (6.5.6))

$$W^{\mu\nu} \equiv \frac{1}{2} \int \mathrm{d}^3 p' \sum_{\lambda_3} \sum_{\lambda_4} J^{+\mu} J^\nu \delta^4 (p + q - p') \tag{I.4.5}$$

where $\sum_{\lambda_3}$ and $\sum_{\lambda_4}$ are the summation for the spin states of the initial and final state particles, and $\frac{1}{2}$ is the spin state averaging factor for the initial proton. We have obtained in Eq. (6.5.13) the expression of the hadronic current

$$\begin{aligned} W^{\mu\nu} = &\left(\frac{q^\mu q^\nu}{q^2} - g^{\mu\nu}\right) W_1(\nu, q^2) \\ &+ \left(p^\mu - \frac{p\cdot q}{q^2} q^\mu\right)\left(p^\nu - \frac{p\cdot q}{q^2} q^\nu\right) \frac{W_2(\nu, q^2)}{M_{\mathrm{N}}^2} \end{aligned} \tag{I.4.6}$$

For the scattering of the unpolarized electron beam on unpolarized proton, averaging for the spin state of the initial particle and summation for the final state particle spin states are needed. Therefore, the cross section is

$$\begin{aligned} \sigma &= \frac{1}{4} \iint \frac{M_p}{2p\cdot k} \sum |A_{\mathrm{NE}}|^2 \frac{\mathrm{d}^3 p'}{(2\pi)^3} \frac{\mathrm{d}^3 k'}{(2\pi)^3} \frac{1}{2E'_k} (2\pi)^4 \delta^4(k + p - k' - p') \\ &= \frac{1}{4} \iint \frac{M_p}{2p\cdot k} \frac{e^4}{q^4} \sum_{\lambda_1\lambda_2\lambda_3\lambda_4} (j_\mu^+ j_\nu J^{\mu+} J^\nu) \frac{\mathrm{d}^3 p'}{(2\pi)^3} \frac{\mathrm{d}^3 k'}{(2\pi)^3} \frac{1}{2E'_k} (2\pi)^4 \\ &\quad \times \delta^4(k + p - k' - p') \\ &= \iint \frac{M_p}{2p\cdot k} \frac{e^4}{q^4} \frac{1}{2} \sum_{\lambda_1\lambda_2} j_\mu^+ j_\nu \frac{1}{2} \sum_{\lambda_3\lambda_4} J^{\mu+} J^\nu \frac{\mathrm{d}^3 p'}{(2\pi)^3} \frac{\mathrm{d}^3 k'}{(2\pi)^3} \frac{1}{2E'_k} (2\pi)^4 \\ &\quad \times \delta^4(k + p - k' - p') \end{aligned}$$

$$= \iint \frac{M_p}{2p\cdot k}\frac{e^4}{q^4}\left(\frac{1}{2}\sum_{\lambda_1\lambda_2} j_\mu^+ j_\nu\right)\left(\frac{1}{2}\mathrm{d}^3p'\sum_{\lambda_3\lambda_4} J^{\mu+}J^\nu\delta^4(k+p-k'-p')\right)$$

$$\times \frac{\mathrm{d}^3k'}{(2\pi)^3}\frac{1}{2E_k'}(2\pi)$$

$$= \int \frac{M_p}{2p\cdot k}\frac{e^4}{q^4}L_{\mu\nu}W^{\mu\nu}\frac{\mathrm{d}^3k'}{(2\pi)^3}\frac{1}{2E_k'}(2\pi) \tag{I.4.7}$$

where the factor $\frac{1}{4}$ on the left-hand side of the above equation is the averaging factor for the spin states of the initial particles, and $\sum$ stands for the summation for the initial particle spin states. The phase space factor for the final hadronic state in the inelastic scattering is defined as $\frac{\mathrm{d}^3p'}{(2\pi)^3}$. It can be seen from eq. (I.4.7) that, to calculate the scattering cross section, we need first to calculate the scalar product of the leptonic and hadronic currents $L_{\mu\nu}W^{\mu\nu}$. Using Eqs. (I.4.4) and (I.4.6), we can obtain

$$L_{\mu\nu}W^{\mu\nu} = 2\left(k_\mu k_\nu' + k_\nu k_\mu' + \frac{q^2}{2}g_{\mu\nu}\right)\left[\left(\frac{q^\mu q^\nu}{q^2} - g^{\mu\nu}\right)W_1(\nu,q^2)\right.$$

$$\left. + \left(p^\mu - \frac{p\cdot q}{q^2}q^\mu\right)\left(p^\nu - \frac{p\cdot q}{q^2}q^\nu\right)\frac{W_2(\nu,q^2)}{M_{\mathrm{N}}^2}\right] \tag{I.4.8}$$

Using the scalar products of momenta given in Eq. (I.3.20), Eq. (I.4.8) can be simplified to be

$$L_{\mu\nu}W^{\mu\nu} = 8E_kE_k'\sin^2\frac{\theta}{2}W_1(\nu,q^2) + 4E_kE_k'\cos^2\frac{\theta}{2}W_2(\nu,q^2)$$

$$= 4E_kE_k'\left(2\sin^2\frac{\theta}{2}W_1(\nu,q^2) + \cos^2\frac{\theta}{2}W_2(\nu,q^2)\right) \tag{I.4.9}$$

Substituting Eq. (I.4.9) into Eq. (I.4.7), we obtain

$$\sigma = \int \frac{M_p}{2p\cdot k}\frac{e^4}{q^4}\frac{\mathrm{d}^3k'}{(2\pi)^3}\frac{1}{2E_k'}(2\pi)$$

$$\cdot 4E_kE_k'(2\sin^2\frac{\theta}{2}W_1(\nu,q^2) + \cos^2\frac{\theta}{2}W_2(\nu,q^2))$$

$$= \int \frac{1}{2E_k}\frac{e^4}{q^4}\frac{k'^2\mathrm{d}k'\mathrm{d}\Omega}{(2\pi)^2}\frac{1}{2E_k'} \times 4E_kE_k'\left(2\sin^2\frac{\theta}{2}W_1(\nu,q^2)\right.$$

$$\left. + \cos^2\frac{\theta}{2}W_2(\nu,q^2)\right)$$

$$= \int \frac{e^4}{q^4}\frac{k'^2\mathrm{d}k'\mathrm{d}\Omega}{(2\pi)^2}\left(2\sin^2\frac{\theta}{2}W_1(\nu,q^2) + \cos^2\frac{\theta}{2}W_2(\nu,q^2)\right) \tag{I.4.10}$$

For the case of neglecting the mass of electron

$$k' = |\vec{k'}| = E'_k, \quad \mathrm{d}k' = \mathrm{d}E'_k \tag{I.4.11}$$

Equation (I.4.10) can be changed to

$$\sigma = \int \frac{e^4}{q^4}\frac{E'^2_k \mathrm{d}E'_k \mathrm{d}\Omega}{(2\pi)^2}\left(2\sin^2\frac{\theta}{2}W_1(\nu,q^2) + \cos^2\frac{\theta}{2}W_2(\nu,q^2)\right) \tag{I.4.12}$$

Substituting $q^2 = -4E_kE'_k\sin^2\frac{\theta}{2}$ in Eq. (I.3.20) into the above equation and after a few steps of simplification, we can obtain

$$\sigma = \int \mathrm{d}E'_k \mathrm{d}\Omega \frac{\alpha^2}{4E^2_k \sin^4\dfrac{\theta}{2}}\left(2\sin^2\frac{\theta}{2}W_1(\nu,q^2) + \cos^2\frac{\theta}{2}W_2(\nu,q^2)\right) \tag{I.4.13a}$$

where $\alpha = e^2/(4\pi)$ is used. If change the symbol of the electron energy in the final state $E_k \to E$, $E'_k \to E'$, Eq. (I.4.13a) can be written as

$$\sigma = \int \mathrm{d}E' \mathrm{d}\Omega \frac{\alpha^2}{4E^2 \sin^4\frac{\theta}{2}}\left(2\sin^2\frac{\theta}{2}W_1(\nu,q^2) + \cos^2\frac{\theta}{2}W_2(\nu,q^2)\right) \tag{I.4.13b}$$

It can be seen from the above equation that the differential cross section is

$$\frac{\sigma}{\mathrm{d}E'\mathrm{d}\Omega} = \frac{\alpha^2}{4E^2\sin^4\frac{\theta}{2}}\left(2\sin^2\frac{\theta}{2}W_1(\nu,q^2) + \cos^2\frac{\theta}{2}W_2(\nu,q^2)\right) \tag{I.4.14}$$

I.5. Calculation of the Differential Cross Section of Electron–Parton Scattering

Assume that partons in nucleon are fermions of spin $\frac{1}{2}$, charge Q_i, mass m_i, and the coupling vertex of the parton and electromagnetic field is $\mathrm{i}Q_ie\gamma_\mu$. The calculation of the cross section of electron–parton scattering is the same as that of scattering of electron and a spin-$\frac{1}{2}$ point-like fermion given in Appendix I.1, where only the mass, charge and the fermion momenta need to be replaced with the relevant mass, charge and momenta of parton. Therefore, we can use the result in Eq. (I.1.10), and change the charge of proton $e \to Q_ie$, the momenta $p \to p_i$, $p' \to p'_i$, and mass $M_p \to m_i$, where p_i and p'_i are the four-momenta of the initial and final state partons, respectively. With these replacement, we can immediately obtain the scattering cross section of electron and patron

$$\sigma = \int \frac{Q_i^2 e^4}{q^4} \frac{1}{p_i \cdot k}(p_i \cdot k p_i' \cdot k' + p_i \cdot k' \ p_i' \cdot k - m_i^2 k \cdot k') \\ \cdot \frac{E_k' \mathrm{d}E_k' \mathrm{d}\Omega}{(2\pi)^2} \delta((k + p_i - k')^2 - m_i^2) \tag{I.5.1}$$

Using $q = k - k'$, the definition $\nu = \frac{p \cdot q}{M_\mathrm{N}}$ and the relations in the infinite momentum frame

$$p_i = x_i p, \quad m_i \approx x_i M_\mathrm{N} \tag{I.5.2}$$

the δ function in Eq. (I.5.1) can be changed to

$$\begin{aligned} \delta((k + p_i - k')^2 - m_i^2) &= \delta((p_i + q)^2 - m_i^2) \\ &= \delta(2p_i \cdot q + q^2) \\ &= \delta(2x_i p \cdot q + q^2) \\ &= \delta(2x_i M_\mathrm{N} \nu - Q^2) \\ &= \frac{1}{2M_\mathrm{N}\nu} \delta\left(x_i - \frac{Q^2}{2M_\mathrm{N}\nu}\right) \end{aligned} \tag{I.5.3}$$

Substituting Eq. (I.5.3) into Eq. (I.5.1), we can obtain

$$\sigma = \int \frac{Q_i^2 e^4}{q^4} \frac{1}{p_i \cdot k}(p_i \cdot k p_i' \cdot k' + p_i \cdot k' p_i' \cdot k - m_i^2 k \cdot k') \\ \cdot \frac{E_k' \mathrm{d}E_k' \mathrm{d}\Omega}{2(2\pi)^2 M_\mathrm{N}\nu} \delta\left(x_i - \frac{Q^2}{2M_\mathrm{N}\nu}\right) \tag{I.5.4}$$

In the infinite momentum frame, the momentum scalar products in the above equation are

$$\begin{cases} p_i \cdot k = x_i p \cdot k \\ p_i \cdot k' = x_i p \cdot k' \\ p_i' \cdot k = (q + p_i) \cdot k = q \cdot k + x_i p \cdot k = \dfrac{1}{2} q^2 + x_i p \cdot k \\ p_i' \cdot k' = (q + p_i) \cdot k' = q \cdot k' + x_i p \cdot k' = -\dfrac{1}{2} q^2 + x_i p \cdot k' \\ k \cdot k' = -\dfrac{1}{2} q^2 \end{cases} \tag{I.5.5}$$

The scalar products $p \cdot k$ and $p \cdot k'$ in Eq. (I.5.5) are all Lorentz invariant quantities, so their results can be expressed in the laboratory frame, then

Eq. (I.5.5) becomes

$$\begin{cases} p_i \cdot k = x_i M_{\mathrm{N}} E_k \\ p_i \cdot k' = x_i M_{\mathrm{N}} E_k' \\ p_i' \cdot k = (q+p_i)\cdot k = q\cdot k + x_i p \cdot k = \dfrac{1}{2}q^2 + x_i M_{\mathrm{N}} E_k \\ p_i' \cdot k' = (q+p_i)\cdot k' = q\cdot k' + x_i p\cdot k' = -\dfrac{1}{2}q^2 + x_i M_{\mathrm{N}} E_k' \\ k\cdot k' = -\dfrac{1}{2}q^2 \end{cases} \tag{I.5.6}$$

In addition, according to

$$\begin{cases} q^2 = (k-k')^2 = -2k\cdot k' = -2E_k E_k'(1-\cos\theta) = -4E_k E_k' \sin^2\dfrac{\theta}{2} \\ (q+p_i)^2 = p_i'^2 \Rightarrow q^2 = -2q\cdot p_i = -2x_i q\cdot p = -2x_i M_{\mathrm{N}}(E_k - E_k') \end{cases} \tag{I.5.7}$$

we can obtain

$$\begin{cases} E_k' = \dfrac{E_k}{1+\frac{2E_k}{x_i M_{\mathrm{N}}}\sin^2\frac{\theta}{2}} \\ q^2 = -\dfrac{4E_k^2 \sin^2\dfrac{\theta}{2}}{1+\frac{2E_k}{x_i M_{\mathrm{N}}}\sin^2\frac{\theta}{2}} \end{cases} \tag{I.5.8}$$

Substituting Eqs. (I.5.6) and (I.5.8) and $m_i = x_i M_{\mathrm{N}}$ into Eq. (I.5.4), and simplifying it, we get

$$\sigma = \int \mathrm{d}E_k'\mathrm{d}\Omega \frac{Q_i^2\alpha^2}{4E_k^2\sin^4\dfrac{\theta}{2}}\frac{x_i}{\nu}\left(\cos^2\frac{\theta}{2} + \frac{1}{x_i^2 M_{\mathrm{N}}^2}\frac{2E_k^2\sin^4\frac{\theta}{2}}{1+\frac{2E_k}{x_i M_{\mathrm{N}}}\sin^2\frac{\theta}{2}}\right) \times \delta\left(x_i - \frac{Q^2}{2M_{\mathrm{N}}\nu}\right) \tag{I.5.9}$$

With the result of q^2 given in Eq. (I.5.8), and the definition $Q^2 = -q^2$, Eq. (I.5.9) can be expressed more compactly

$$\sigma = \int \mathrm{d}E_k'\mathrm{d}\Omega \frac{Q_i^2\alpha^2}{4E_k^2\sin^4\frac{\theta}{2}}\frac{x_i}{\nu}\left(\cos^2\frac{\theta}{2} + \frac{Q^2}{2x_i^2 M_{\mathrm{N}}^2}\sin^2\frac{\theta}{2}\right)\delta\left(x_i - \frac{Q^2}{2M_{\mathrm{N}}\nu}\right) \tag{I.5.10}$$

From the above equation, we can obtain the differential cross section of the electron–parton scattering

$$\frac{\sigma}{\mathrm{d}E_k'\mathrm{d}\Omega}=\frac{Q_i^2\alpha^2}{4E_k^2\sin^4\dfrac{\theta}{2}}\frac{x_i}{\nu}\left(\cos^2\frac{\theta}{2}+\frac{Q^2}{2x_i^2M_\mathrm{N}^2}\sin^2\frac{\theta}{2}\right)\delta\left(x_i-\frac{Q^2}{2M_\mathrm{N}\nu}\right) \tag{I.5.11}$$

If the initial and final state energies of electrons are expressed as E and E', respectively, then we have

$$\frac{\sigma}{\mathrm{d}E'\mathrm{d}\Omega}=\frac{Q_i^2\alpha^2}{4E^2\sin^4\frac{\theta}{2}}\frac{x_i}{\nu}\left(\cos^2\frac{\theta}{2}+\frac{Q^2}{2x_i^2M_\mathrm{N}^2}\sin^2\frac{\theta}{2}\right)\delta\left(x_i-\frac{Q^2}{2M_\mathrm{N}\nu}\right) \tag{I.5.12}$$

I.6. Calculation of the Cross Section of Neutrino–Lepton Scattering in the Theory of Four-Fermion Point-Interaction

The early weak interaction theory of Fermi was based on the interaction of four-fermion point-interaction, which contains the characteristic that weak interaction is a short-distance interaction. It neglects the details of the weak interaction in the microscale. The theory of point-interaction achieved great success in describing the experiment of weak interaction. However, with the increase of momentum transfer in the process of weak interaction, the microscopic details of the interaction will gradually emerge. This would show the limits of point-interaction. We have discussed this problem in Sec. 7.8. Now we give the calculation of the cross section of neutrino–lepton scattering process $\nu_\mu+\mathrm{e}^-\to\mu^-+\nu_\mathrm{e}$ under the theory of four-fermion point-interaction. In the (V–A) current–current interaction theory, the Hamiltonian describing $\nu_\mu+\mathrm{e}^-\to\mu^-+\nu_\mathrm{e}$ scattering process is

$$H_\mathrm{W}=\frac{G_\mu}{\sqrt{2}}\bar{\nu}_\mathrm{e}\gamma_\mu\left(1-\gamma_5\right)e\bar{\mu}\gamma^\mu(1-\gamma_5)\nu_\mu \tag{I.6.1}$$

The resulting interaction vertex of $(\mathrm{e}\nu_\mathrm{e})(\mu\nu_\mu)$ is $\frac{G_\mu}{\sqrt{2}}\gamma_\mu(1-\gamma_5)\gamma^\mu(1-\gamma_5)$. The Feynman diagram for $\nu_\mu+\mathrm{e}^-\to\mu^-+\nu_\mathrm{e}$ scattering process is shown in Fig. I.1, where k_1, k_2, k_3 and k_4 are the momenta of ν_μ, e, μ and ν_e,

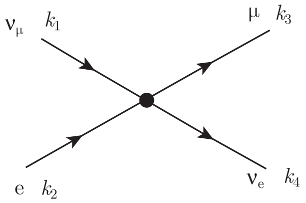


Fig. I.1.

respectively. The amplitude of this scattering process is

$$A = \frac{G_{\mathrm{F}}}{\sqrt{2}} \bar{u}(k_3)\gamma_\mu(1-\gamma_5)u(k_1)\bar{u}(k_4)\gamma^\mu(1-\gamma_5)u(k_2) \tag{I.6.2}$$

where the spin indices are omitted.

Calculate the modular square of the scattering amplitude and sum up the spin state of all particles

$$\begin{aligned}
\sum_\lambda |A|^2 &= \frac{G_{\mathrm{F}}^2}{2}[\bar{u}(k_3)\gamma_\mu(1-\gamma_5)u(k_1)\bar{u}(k_4)\gamma^\mu(1-\gamma_5)u(k_2)]^+ \\
&\quad \cdot[\bar{u}(k_3)\gamma_\nu(1-\gamma_5)u(k_1)\bar{u}(k_4)\gamma^\nu(1-\gamma_5)u(k_2)] \\
&= \frac{G_{\mathrm{F}}^2}{2}[\bar{u}(k_1)\gamma_\mu(1-\gamma_5)u(k_3)\bar{u}(k_2)\gamma^\mu(1-\gamma_5)u(k_4)] \\
&\quad \cdot[\bar{u}(k_3)\gamma_\nu(1-\gamma_5)u(k_1)\bar{u}(k_4)\gamma^\nu(1-\gamma_5)u(k_2)] \\
&= \frac{G_{\mathrm{F}}^2}{2}[\bar{u}(k_1)\gamma_\mu(1-\gamma_5)u(k_3)\bar{u}(k_3)\gamma_\nu(1-\gamma_5)u(k_1)] \\
&\quad \cdot[\bar{u}(k_2)\gamma^\mu(1-\gamma_5)u(k_4)\bar{u}(k_4)\gamma^\nu(1-\gamma_5)u(k_2)] \\
&= \frac{G_{\mathrm{F}}^2}{2}\mathrm{Tr}[u(k_1)\bar{u}(k_1)\gamma_\mu(1-\gamma_5)u(k_3)\bar{u}(k_3)\gamma_\nu(1-\gamma_5)] \\
&\quad \cdot\mathrm{Tr}[u(k_2)\bar{u}(k_2)\gamma^\mu(1-\gamma_5)u(k_4)\bar{u}(k_4)\gamma^\nu(1-\gamma_5)] \\
&= \frac{G_{\mathrm{F}}^2}{2}\mathrm{Tr}[\not{k}_1\gamma_\mu(1-\gamma_5)\ \not{k}_3\gamma_\nu(1-\gamma_5)]\mathrm{Tr} \\
&\quad \times[\not{k}_2\gamma^\mu(1-\gamma_5)\ \not{k}_4\gamma^\nu(1-\gamma_5)] \\
&= 2G_{\mathrm{F}}^2\mathrm{Tr}[\not{k}_1\gamma_\mu\ \not{k}_3\gamma_\nu(1-\gamma_5)]\mathrm{Tr}[\not{k}_2\gamma^\mu\ \not{k}_4\gamma^\nu(1-\gamma_5)]
\end{aligned} \tag{I.6.3}$$

Calculating the traces in the above equation, we get

$$\sum_\lambda |A|^2 = 128G_{\mathrm{F}}^2 k_1\cdot k_2 k_3\cdot k_4 \tag{I.6.4}$$

The scattering process is a massless fermion scattering process when all lepton masses are ignored. From Eq. (3.4.13), the cross section of this process is

$$\sigma = \iint \frac{1}{4k_1 \cdot k_2} \sum_{\lambda} |A|^2 \frac{\mathrm{d}^3 k_3}{(2\pi)^3} \frac{1}{2E_3} \frac{\mathrm{d}^3 k_4}{(2\pi)^3} \frac{1}{2E_4} (2\pi)^4 \delta^4(k_1 + k_2 - k_3 - k_4) \tag{I.6.5}$$

In the calculation of this cross section, all the spin states of the initial and final state particles are summed, without considering the averaging factor for the initial state particles, because only the left-handed current contributes to the scattering amplitude in the V–A theory, so only the left-handed fermion contributes to the scattering process, while right-handed neutrinos and right-handed electrons do not contribute.

Substituting the result of Eq. (I.6.4) into Eq. (I.6.5), one can obtain

$$\begin{aligned} \sigma &= \iint 32G_{\mathrm{F}}^2 k_3 \cdot k_4 \frac{\mathrm{d}^3 k_3}{(2\pi)^3} \frac{1}{2E_3} \frac{\mathrm{d}^3 k_4}{(2\pi)^3} \frac{1}{2E_4} (2\pi)^4 \delta^4(k_1 + k_2 - k_3 - k_4) \\ &= \iint 32G_{\mathrm{F}}^2 k_3 \cdot k_4 \frac{\mathrm{d}^3 k_3}{(2\pi)^3} \frac{1}{2E_3} \frac{\mathrm{d}^4 k_4}{(2\pi)^3} \delta(k_4^2) (2\pi)^4 \delta^4(k_1 + k_2 - k_3 - k_4) \\ &= \int 32G_{\mathrm{F}}^2 k_3 \cdot (k_1 + k_2 - k_3) \frac{\mathrm{d}^3 k_3}{(2\pi)^2} \frac{1}{2E_3} \delta((k_1 + k_2 - k_3)^2) \end{aligned} \tag{I.6.6}$$

In the center-of-mass coordinate frame, we have

$$k_1 = (E, \vec{k}), \quad k_2 = (E, -\vec{k}), \quad k_1 + k_2 = (2E, 0) \tag{I.6.7}$$

where $\vec{k}$ is the 3-dimensional momentum of the initial particle in the center-of-mass frame. Since the masses of leptons are dropped, so we have

$$k_1^2 = k_2^2 = E^2 - k^2 = 0, \quad k_3^2 = 0 \tag{I.6.8}$$

where $k \equiv |\vec{k}|$. From Eqs. (I.6.7) and (I.6.8), we can obtain

$$k_3 \cdot (k_1 + k_2 - k_3) = k_3 \cdot (k_1 + k_2) = 2EE_3 \tag{I.6.9}$$

$$(k_1 + k_2 - k_3)^2 = 4E^2 - 4EE_3 \tag{I.6.10}$$

Substituting Eqs. (I.6.9) and (I.6.10) into Eq. (I.6.6), one obtains

$$\begin{aligned}
\sigma &= \int 32G_{\mathrm{F}}^2 2EE_3 \frac{\mathrm{d}^3 k_3}{(2\pi)^2} \frac{1}{2E_3} \delta(4E^2 - 4EE_3) \\
&= \int 32G_{\mathrm{F}}^2 2EE_3 \frac{k_3^2 \mathrm{d}k_3 \mathrm{d}\Omega}{(2\pi)^2} \frac{1}{2E_3} \delta(4E^2 - 4EE_3) \\
&= \int 32G_{\mathrm{F}}^2 2EE_3 \frac{E_3^2 \mathrm{d}E_3 \mathrm{d}\Omega}{(2\pi)^2} \frac{1}{2E_3} \delta(4E^2 - 4EE_3) \\
&= \int 32G_{\mathrm{F}}^2 E \frac{E_3^2 \mathrm{d}E_3 \mathrm{d}\Omega}{(2\pi)^2} \delta(4E^2 - 4EE_3) \\
&= \int 8G_{\mathrm{F}}^2 \frac{E_3^2 \mathrm{d}E_3 \mathrm{d}\Omega}{(2\pi)^2} \delta(E - E_3) \\
&= \frac{8G_{\mathrm{F}}^2 E^2}{\pi} \\
&= \frac{8G_{\mathrm{F}}^2 k^2}{\pi} \qquad \text{(I.6.11)}
\end{aligned}$$

This result indicates that, with the increase of the central-of-mass energy, the cross section of electron–neutrino scattering process increases infinitely in the four-fermion point-interaction theory, which is obviously contrary to the unitary limit of the cross section.

Appendix II

Table of Physical Constants

Table II.1. Physical Constants [1].

Quantity	Symbol	Value
Speed of light in vacuum	c	$299\ 792\ 458$ m s^{-1}
Planck constant	h	$6.626\ 070\ 040(81) \times 10^{-34}$ J s
	$\hbar \equiv h/2\pi$	$1.054\ 571\ 800(13) \times 10^{-34}$ J s $= 6.582\ 119\ 514(40)$ $\times 10^{-22}$ MeV s
Magnitude of electron charge	e	$1.602\ 176\ 620\ 8(98) \times 10^{-19}$ C $= 4.803\ 204\ 673(30) \times 10^{-10}$ esu
	$\hbar c$	$197.326\ 978\ 8(12)$ MeV fm
	$(\hbar c)^2$	$0.389\ 379\ 365\ 6(48)$ GeV^2 mbarn
Mass of electron	m_{e}	$0.510\ 998\ 946\ 1(31)\mathrm{MeV}/c^2$ $= 9.109\ 383\ 56(11) \times 10^{-31}\mathrm{kg}$
Mass of proton	m_{p}	$938.272\ 081\ 3(58)\mathrm{MeV}/c^2$ $= 1.672\ 621\ 898(21) \times 10^{-27}\mathrm{kg}$ $= 1.007\ 276\ 466\ 879(91)\mathrm{u}$ $= 1\ 836.152\ 673\ 89(17) m_{\mathrm{e}}$
Mass of deuteron	m_{d}	$1\ 875.612\ 928(12)$ MeV/c^2
Atomic mass unit (u)	(mass ^{12}C atom)/12 $= (1\ \mathrm{g})/(N_{\mathrm{A}}\ \mathrm{mol})$	$931.494\ 095\ 4(57)\mathrm{MeV}/c^2$ $= 1.660\ 539\ 040(20) \times 10^{-27}\mathrm{kg}$
Free space permittivity	ε_0	$8.854\ 187\ 817 \ldots \times 10^{-12}\mathrm{C}^2$ $\mathrm{N}^{-1}\ \mathrm{m}^{-2}$

(*Continued*)

Table II.1. (*Continued*)

Quantity	Symbol	Value
Free space permeability	μ_0	$4\pi \times 10^{-7}\mathrm{NA}^{-2}$ $= 12.566\ 370\ 614\ldots \times 10^{-7}\,\mathrm{N\,A}^{-2}$
Fine-structure constant	$\alpha = e^2/4\pi\varepsilon_0\hbar c$	$7.297\ 352\ 566\ 4(17) \times 10^{-3}$ $= 1/137.035\ 999\ 139(31)$
Classical electron radius	$r_\mathrm{e} = e^2/4\pi\varepsilon_0 m_e c^2$	$2.817\ 940\ 322\ 7(19) \times 10^{-15}\,\mathrm{m}$
(Electron Compton wavelength)/2π	$\lambda\!\!\!^{-}_\mathrm{e} = \hbar/m_\mathrm{e}c = r_\mathrm{e}\alpha^{-1}$	$3.861\ 592\ 676\ 4(18) \times 10^{-13}\,\mathrm{m}$
Bohr radius ($m_p \to \infty$)	$a_\infty = 4\pi\varepsilon_0\hbar^2/m_\mathrm{e}e^2$ $= r_\mathrm{e}\alpha^{-2}$	$0.529\ 177\ 210\ 67(12) \times 10^{-10}\ \mathrm{m}$
Thomson cross section	$\sigma_\mathrm{T} = 8\pi r_\mathrm{e}^2/3$	$0.665\ 245\ 871\ 58(91)$ barn
Bohr magneton	$\mu_\mathrm{B} = e\hbar/2m_\mathrm{e}$	$5.788\ 381\ 801\ 2(26)\times 10^{-11}\ \mathrm{MeV\ T}^{-1}$
Nuclear magneton	$\mu_\mathrm{N} = e\hbar/2m_\mathrm{p}$	$3.152\ 451\ 255\ 0(15) \times 10^{-14}\,\mathrm{MeV\ T}^{-1}$
Newtonian Gravitational constant	G_N	$6.674\ 08(31)\times 10^{-11}\ \mathrm{m}^3\ \mathrm{kg}^{-1}\ \mathrm{s}^{-2}$ $= 6.708\ 61(31) \times 10^{-39}\hbar c(\mathrm{GeV}/c^2)^{-2}$
Standard gravitational acceleration	g_N	$9.806\ 65\,\mathrm{m\,s}^{-2}$
Avogadro constant	N_A	$6.022\ 140\ 857(74) \times 10^{23}\ \mathrm{mol}^{-1}$
Boltzmann constant	k	$1.380\ 648\ 52(79) \times 10^{-23}\ \mathrm{J\ K}^{-1}$ $= 8.617\ 3303(50) \times 10^{-5}\,\mathrm{eV\ K}^{-1}$
Fermi constant	$G_\mathrm{F}/(\hbar c)^3$	$1.166\ 378\ 7(6) \times 10^{-5}\,\mathrm{GeV}^{-2}$
Weak mixing angle	$\sin^2\hat{\theta}(M_\mathrm{Z})(\overline{\mathrm{MS}})$	$0.231\ 29(5)$
Mass of W$^\pm$ boson	m_W	$80.385(15)\ \mathrm{GeV}/c^2$
Mass of Z^0 boson	M_Z	$91.1876(21)\ \mathrm{GeV}/c^2$
Strong coupling constant	$\alpha_\mathrm{S}(m_\mathrm{Z})$	$0.1182(12)$
$\pi = 3.141\ 592\,653\,589\,793\,238$	$\mathrm{e} = 2.718\ 281\,828\,459\,045\,235$	$\gamma = 0.577\ 215\ 664\ 901\ 532\ 861$
$1\,\mathrm{in} \equiv 0.0254\,\mathrm{m}$	$1\mathrm{G} \equiv 10^{-4}\ \mathrm{T}$	$1\,\mathrm{eV} = 1.602\ 176\ 620\ 8(98) \times 10^{-19}\ \mathrm{J}$
$1\ \text{Å} \equiv 0.1\,\mathrm{nm}$	$1\ \mathrm{dyne} \equiv 10^{-5}\ \mathrm{N}$	$1\,\mathrm{eV}/c^2 = 1.782\ 661\ 907(11) \times 10^{-36}\ \mathrm{kg}$
$0^\circ\mathrm{C} \equiv 273.15\,\mathrm{K}$	$1\ \mathrm{barn} \equiv 10^{-28}\ \mathrm{m}^2$	$1\,\mathrm{erg} \equiv 10^{-7}\ \mathrm{J}$
$2.997\ 924\ 58 \times 10^9\mathrm{esu} = 1\mathrm{C}$	$1\mathrm{atmosphere} = 760\,\mathrm{Torr}$ $\equiv 101\ 325\,\mathrm{Pa}$	

Table II.2. Astrophysical Constants.

Quantity	Symbol	Value	Reference
Newtonian gravitational constant	$G_{\rm N}$	$6.674\ 08(31) \times 10^{-11}\,{\rm m}^3\,{\rm kg}^{-1}\,{\rm s}^{-2}$	[2]
Planck mass	$\sqrt{\hbar c/G_{\rm N}}$	$1.220\ 910(29) \times 10^{19}\ {\rm GeV}/c^2$ $= 2.176\ 47(5) \times 10^{-8}\,{\rm kg}$	[2]
Planck length	$\sqrt{\hbar G_{\rm N}/c^3}$	$1.616\ 229(38) \times 10^{-35}\,{\rm m}$	[2]
Standard gravitational acceleration	$g_{\rm N}$	$9.806\ 65{\rm m\,s}^{-2}$	[2]
Jansky (flux density)	$\rm J_y$	$10^{-26}\,{\rm W\,m}^{-2}\,{\rm Hz}^{-1}$	Definition
Tropical year	yr	$31\ 556\ 925.2\ {\rm s} \approx \pi \times 10^7\,{\rm s}$	[3]
Sidereal year		$31\ 558\ 149.8\ {\rm s} \approx \pi \times 10^7\,{\rm s}$	[3]
Mean sidereal day		$23^{\rm h}\ 56^{\rm m}\ 04.^{\rm s}090\ 53$	[3]
Astronomical unit	au	149 597 870 700 m	[1]
parsec (1 au/1 arc sec)	pc	$3.085\ 677\ 581\ 49 \times 10^{16}\,{\rm m}$ $= 3.262 \ldots$ ly	[1]
Light year	ly	$0.306\ 6 \ldots {\rm pc} = 0.946\ 053 \ldots \times 10^{16}\,{\rm m}$	[1]
Schwarzschild radius of the Sun	$2G_{\rm N}M_\odot/c^2$	$2.953\ 250\ 24\,{\rm km}$	[1]
Solar mass	$M_\odot$	$1.988\ 48(9) \times 10^{30}\,{\rm kg}$	[1]
Nominal Solar equatorial radius	$R_\odot$	$6.957 \times 10^8\,{\rm m}$	[1]
Nominal Solar constant	$S_\odot$	$1\ 361\,{\rm W\,m}^{-2}$	[1]
Nominal Solar photosphere temperature	$T_\odot$	$5\ 772\,{\rm K}$	[1]
Nominal Solar luminosity	$L_\odot$	$3.828 \times 10^{26}\,{\rm W}$	[1]
Schwarzschild radius of the Earth	$2G_{\rm N}M_\oplus/c^2$	$8.870\ 056\ 580(18)\,{\rm mm}$	[1]
Mass of the earth	$M_\oplus$	$5.972\ 4(3) \times 10^{24}\,{\rm kg}$	[1]
Nominal Earth equatorial radius	$R_\oplus$	$6.3781 \times 10^6\,{\rm m}$	[1]
Luminosity conversion	L	$3.0128 \times 10^{28} \times 10^{-0.4\ {\rm M_{bol}}}\ {\rm W}$	[4]
Flux conversion	F	$2.5180 \times 10^{-8} \times 10^{-0.4{\rm m_{bol}}}\ {\rm W\,m}^{-2}$	[5]

(*Continued*)

Table II.2. (*Continued*)

Quantity	Symbol	Value	Reference
Absolute monochromatic magnitude	AB	$-2.5\log_{10} f_\nu - 56.10$ (f_ν in $\mathrm{Wm^{-2}\ Hz^{-1}}$) $= -2.5\log_{10} f_\nu + 8.90$($f_\nu$ in Jy)	[1]
Solar angular velocity around the Galactic center	Θ_0/R_0	$30.3 \pm 0.9\,\mathrm{km\,s^{-1}\,kpc^{-1}}$	[1]
Solar distance from Galactic center	R_0	$8.00 \pm 0.25\,\mathrm{kpc}$	[1]
Circular velocity at R_0	v_0 or Θ_0	$254(16)\,\mathrm{km\,s^{-1}}$	[1]
Escape velocity from Galaxy	v_{esc}	$498\,\mathrm{km/s} < v_{\mathrm{esc}} < 608\,\mathrm{km/s}$	[6]
Local disk density	ρ_{disk}	$(3\sim12) \times 10^{-24}\,\mathrm{g\,cm^{-3}} \approx 2\sim7\,\mathrm{GeV}/c^2\,\mathrm{cm^{-3}}$	[7]
Local dark matter density	ρ_χ	$0.3\,\mathrm{GeV}/c^2\,\mathrm{cm^{-3}}$ (within factor $2 \sim 3$)	[1]
Present-day CMB temperature	T_0	$2.725\ 5(6)\,\mathrm{K}$	[8, 9]
Present day CMB dipole amplitude	d	$3.364\ 5(20)\,\mathrm{mK}$	[8, 10]
Solar velocity with respect to CMB	$v_\odot$	$369(1)\,\mathrm{km\,s^{-1}}$ towards $(\ell, b) = (263.99(14)^\circ\ , 48.26(3)^\circ)$	[8, 10]
Local Group velocity with respect to CMB	v_{LG}	$627(22)\,\mathrm{km\,s^{-1}}$ towards $(\ell, b) = (276(3)^\circ, 30(3)^\circ)$	[8, 10]
Number density of CMB photons	n_γ	$410.7(T/2.725\ 5)^3\,\mathrm{cm^{-3}}$	[1]
Density of CMB photons	ρ_γ	$4.645(4)(T/2.725\ 5)^4$ $\times 10^{-34}\mathrm{g\ cm^{-3}}$ $\approx 0.260\,\mathrm{eVcm^{-3}}$	[1]
Entropy density/Boltzmann constant	s/k	$2\ 891.2\ (T/2.725\ 5)^3\,\mathrm{cm^{-3}}$	[1]
Present day Hubble expansion rate	H_0	$100\,h\,\mathrm{km\ s^{-1}\,Mpc^{-1}} = h\times (9.777\ 752\ \mathrm{Gyr})^{-1}$	[1]

Table II.2. (*Continued*)

Quantity	Symbol	Value	Reference
Scale factor for Hubble expansion rate	h	$0.678(9)$	[10, 11]
Hubble length	c/H_0	$0.925\,062\,9 \times 10^{26}\,h^{-1}\,\mathrm{m}$ $= 1.374(18) \times 10^{26}\,\mathrm{m}$	
Scale factor for cosmological constant	$c^2/3H_0^2$	$2.852\,47 \times 10^{51} h^{-2}\,\mathrm{m}^2$ $= 6.20(17) \times 10^{51}\,\mathrm{m}^2$	
Critical density of the Universe	$\rho_{\mathrm{crit}} = 3H_0^2/8\pi G_{\mathrm{N}}$	$1.87840(9) \times 10^{-29} h^2\,\mathrm{g\,cm^{-3}}$ $= 1.053\,71(5)$ $\times 10^{-5} h^2\,(\mathrm{GeV}/c^2)\,\mathrm{cm}^{-3}$	
Ratio of baryon-to-photon number density	$\eta = n_{\mathrm{b}}/n_\gamma$	$5.8 \times 10^{-10} \leq \eta \leq 6.6 \times 10^{-10}$ (95% CL)	[12]
Baryon number density	n_{b}	$2.503(26) \times 10^{-7}\,\mathrm{cm}^{-3}$	[13]
CMB radiation density of the Universe	$\Omega_\gamma = \rho_\gamma/\rho_{\mathrm{crit}}$	$2.473 \times 10^{-5} (T/2.725\,5)^4 h^{-2} = 5.38(15) \times 10^{-5}$	[1]
Planck 2015 6-parameter fit to flat ΛCDM cosmology [1]			
Baryon density of the Universe	$\Omega_{\mathrm{b}} = \rho_{\mathrm{b}}/\rho_{\mathrm{crit}}$	$0.02\,226(23)\ h^{-2} = 0.048\,4(10)$	
Cold dark matter density of the Universe	$\Omega_{\mathrm{CDM}} = \rho_{\mathrm{CDM}}/\rho_{\mathrm{cri}}$	$0.118\,6(20)\ h^{-2} = 0.258(11)$	
100×approx to r_*/D_{A}	$100 \times \theta_{\mathrm{MC}}$	$1.041\,0(5)$	
Reionization optical depth	τ	$0.066(16)$	
Scalar spectral index	n_{s}	$0.968(6)$	
	$\ln(10^{10}\Delta_{\mathrm{R}}^2)$	$3.062(29)$	
Dark energy density of the Universe in ΛCDM model	Ω_Λ	0.692 ± 0.012	
Pressureless matter density of the Universe	$\Omega_{\mathrm{m}} = \Omega_{\mathrm{CDM}} + \Omega_{\mathrm{b}}$	0.308 ± 0.012	

(*Continued*)

Table II.2. (*Continued*)

Quantity	Symbol	Value	Reference
Fluctuation amplitude at $8h^{-1}$ Mpc scale	σ_8	0.815 ± 0.009	
Redshift of matter–radiation equality	z_{eq}	$3\,365 \pm 44$	
Redshift as optical depth equals unity	z_*	$1\,089.9 \pm 0.4$	
Comoving size of sound horizon at z_*	r_*	$144.9 \pm 0.4\,\mathrm{Mpc}$ (Planck CMB)	[14]
Age when optical depth equals unity	t_*	373 kyr	[15]
Redshift at half reionization	z_{reion}	$8.8^{+1.7}_{-1.4}$	
Redshift when acceleration was zero	z_{q}	~ 0.65	[15]
Age of the Universe	t_0	13.80 ± 0.04 Gyr	
Effective number of neutrinos	N_{eff}	3.1 ± 0.6	
Sum of neutrino masses	Σm_ν	$< 0.68\,\mathrm{eV}$ (Planck CMB); $\geq 0.05\,\mathrm{eV}$ (mixing)	
Neutrino density of the Universe	$\Omega_\nu = h^{-2}$ $\Sigma m_{\nu j}/93.04\mathrm{eV}$	< 0.016 (Planck CMB); $\geq 0.001\,2$(mixing)	
Curvature	Ω_{K}	$-0.005^{+0.016}_{-0.017}$(95% CL)	
Running spectral, index slope $k_0 = 0.002\,\mathrm{Mpc}^{-1}$	$\mathrm{d}n_{\mathrm{s}}/\mathrm{dln}k$	$-0.003(15)$	
Tensor-to-scalar field perturbations ratio, $k_0 = 0.002\,\mathrm{Mpc}^{-1}$	$r_{0.002}$ $= T/S$	< 0.114 at 95% CL	
State parameter of dark energy equation	w	-0.97 ± 0.05	[16]
Primordial helium fraction	Y_{p}	0.245 ± 0.004	[17]

References

1. Patrignani C, *et al.* Particle Data Group. *Chinese Phys.* C, 2016, 40: 100001.
2. CODATA recommended 2014 values of the fundamental physical constants: physics.nist.gov/constants.
3. The Astronomical Almanac Online for the year 2016, asa.usno.navy.mil/SecK/Constants.html.
4. M_{bol} = absolute bolometric magnitude = bolometric magnitude at 10 pc.
5. m_{bol} = apparent bolometric magnitude.
6. Smith M C, *et al. Mon. Not. R. Astr. Soc.*, 2007, 379: 755.
7. Gilmore G, Wyse R F G, Kuijken K. *Ann. Rev. Astron. Astrophys*, 1989, 27: 555.
8. Scott D, Smoot G F. *Cosmic Microwave, Background*, Sec. 28 in Patrignani C, *et al.* Particle Data Group. *Chinese Phys.* C, 2016, 40: 100001.
9. Fixsen D. *Astrophys. J.*, 2009, 707: 916.
10. Planck Collab. *2015 Results XIII, Astron. Astrophys.* submitted, arXiv: 1502.01589v2.
11. Lahav O, Liddle A R. *The Cosmological Parameters.* Sec. 25 in Patrignani C, *et al.* Particle Data Group. *Chinese Phys.* C, 2016, 40: 100001.
12. Fields B D, Molarto P, Sarkar S. *Big-Bang Nucleosynthesis* in Patrignani C, *et al.* Particle Data Group. *Chinese Phys.* C, 2016, 40: 100001.
13. Steigman G. *JCAP*, 2006, 10: 016.
14. Weinberg D H, White M. *Dark Energy.* Sec. 27 in Patrignani C, *et al.* Particle Data Group. *Chinese Phys.* C, 2016, 40: 100001.
15. Scott D, Narimani A, Page D N. arXiv: 1309.2381v2.
16. Auborg E, *et al. Phys. Rev.*, 2015, D92: 123516.
17. Aver E, *et al. JCAP*, 2015, 07: 011.

Index

abundances of light elements, 531
Adler anomaly, 491
age of the Universe, 540
angular power spectrum, 569, 571
angular scales, 569
anisotropy, 56, 565, 570
anomaly cancellation, 492
anti-unitary operators, 207
antiscreening, 141
asymptotic freedom, 315, 316
axion, 450, 561

B meson factory, 230
baryon number conservation, 52, 524
baryon–antibaryon asymmetry, 485, 489
baryons, 4, 47
Bhabha scattering, 144, 154
Big Bang model, 574
Big Bang nucleosynthesis, 551
Bjorken's scaling, 316
black-body radiation, 544, 565
bottomonium, 141, 432, 433
box diagram, 228, 239
brane theory, 526

Cabibbo theory, 189, 261
Cabibbo–Kobayashi–Maskawa (CKM) matrix, 181, 286
charge conjugate transformation, 14, 439
charmonium, 432, 433
chiral superfield, 516
chiral symmetry, 203, 445
chiral symmetry breaking, 201
chiral transformation, 445, 448
closed Universe, 532, 534
color confinement, 433
color degree of freedom, 122
comoving frame, 536
comoving observer, 533
cosmic microwave background, 543, 565
cosmic neutrino background, 548
cosmic time, 533
cosmological proper time, 533
covariant derivative, 255
CPT theorem, 33
critical energy density, 534

Dalitz plot, 223
dark energy, 561
dark matter, 571, 572
decay width, 67, 188
decouple, 455, 548
deep inelastic scattering, 144, 167
density parameter, 535
dimensional regularization, 360
dipenguin diagram, 239, 240
dipole, 569, 570
Dirac neutrino, 472
direct CP violation, 212

discrete symmetry, 9, 510
dual tensor, 439

effective Lagrangian, 342
elastic scattering, 58, 60
electric form factor, 162, 164
electromagnetic form factors, 156
electroweak unified theory, 489, 506
entropy density, 549
equation of state, 533, 538
exotic state, 137, 437

Feynman rules, 148
flat Universe, 573
flatness problem, 573
flavor unification, 509
flavor wave functions, 111, 114
four fermion interaction, 197, 198
Friedmann equation, 532
furry theorem, 22

G parity, 93
generation functional, 331, 345
generators, 95, 101, 107
ghost field, 298, 336
GIM mechanism, 180
global gauge symmetry, 10
global supersymmetry, 514
global supersymmetry (supergravity), 514
glueball, 434
gluon, 356
grand unified theory, 490
gravitational lensing, 559, 560
Green function, 321, 328

hadrons, 4
helicity, 26, 30, 425
helicity suppression, 189
Higgs mechanism, 248
horizon problem, 573
hot Big Bang model, 543
Hubble's law, 531
hybrid state, 434

ideal mixing, 114, 115, 133
identical particles, 65
indirect CP violation, 212
inelastic scattering, 5, 60, 166
inflation, 530, 555
infrared fixed point, 417
instanton effect, 449
intermediate boson, 38, 259
inverted mass hierarchies, 483
isospin, 62, 75

Jarlskog invariant, 467

Kobayashi–Maskawa model, 194
Kurie plot, 185

ΛCDM model, 529
Landau–Yang theorem, 24
lattice gauge theory, 431, 437
lepton, 4
lepton mixing matrix, 456
lepton number conservation, 46
linear potential, 141, 433
local gauge invariance, 146, 247, 320
local gauge symmetry, 10, 125, 147
local supersymmetry, 514
local supersymmetry (supergravity), 514

magnetic form factor, 162
Majorana neutrinos, 474
mass formula, 114, 127
membrane theory, 1
meson, 3, 28
minimum subtraction, 383
mixed CP violation, 212
mixing, 129
mixing parameter, 225
modified minimum subtraction, 384
monopole, 569, 570
monopole problem, 574, 577
multi-quark states, 137
Møller scattering, 144

neutrino oscillation, 182, 455
neutrino oscillation in matter, 468

neutrino oscillation in vacuum, 464
non-luminous halo, 558
nonperturbative QCD, 437
normal hierarchy, 482

off-shell subtraction, 383
on-shell subtraction, 381
open Universe, 532
orbital excitation states, 136
OZI Rule, 134

parity non-conservation, 45
particle lifetime, 67
parton distribution, 424
parton model, 5, 144, 169
path integral quantization, 320, 331
Pauli–Villars regularization, 359
Peccei–Quinn symmetry, 450
perturbative QCD, 437
proper time, 533
proton decay, 490, 508
pure leptonic decays, 182

quantum chromodynamics, 125, 439
quantum electrodynamics, 316
quark confinement, 141
quark model, 3, 100

R_ξ gauge, 305
radial excitation states, 137
reciprocity relation, 32, 33
recombination, 567
redshift, 531
regularization, 356, 359
renormalization, 143, 229, 294
renormalization group equation, 409, 413, 416
resonance states, 57
Robertson–Walker, 532
running coupling constant, 415, 416

scattering cross section, 60, 63
sea quarks, 141
See-Saw mechanism, 476
semi-leptonic decays, 182
space inversion, 190
spontaneous symmetry breaking, 247, 497
strange particle, 48, 53
strange quantum number, 212
structure constant, 38, 107, 153
structure functions, 169
superfield, 513, 515
superspace, 513, 514
superstring theory, 526
supersymmetric models, 510

tensor analysis, 104, 111
time reversal, 14, 25

ultraviolet fixed point, 417
unitarity triangle, 235, 237
unitary bound, 196
unitary gauge, 274
unitary operator, 19, 207

V–A theory, 180, 187
vacuum expectation value, 270, 273, 449
valence quarks, 141
vector current conservation, 13
vector superfield, 515

Ward–Takahashi identity, 380, 394
Weisskopf–Wigner approximation, 206
Weyl spinor, 492, 517

Young tableau, 491
Yukawa couplings, 524